Indian Agricultural
Research Institute, New Delhi.

ANNALS

OF THE

NEW YORK ACADEMY OF SCIENCES

VOLUME XLI

CONTENTS OF VOLUME XLI

THE INTERNAL CONSTITUTION OF THE STARS*

By

Zdeněk Kopal, Robert E. Marshak, Henry Norris Russell,
Harlow Shapley, and Jaakko Tuominen

CONTENTS

(1)

INTRODUCTION TO THE CONFERENCE ON THE INTERNAL CONSTITUTION OF THE STARS

By HARLOW SHAPLEY

From the Harvard College Observatory, Cambridge, Massachusetts

Notwithstanding notable progress in the past few years, especially through the work of Dr. Hans Bethe and his collaborators, we are still some distance from completely satisfactory solutions of the problems of the interior energy-sources for stars. For the supergiant stars, for instance, we must worry either about the time-scale or about the efficiency of the mechanism of energy release as it is now visualized and worked out. There is much yet to be done on the problems of opacity, and extensive calculations are needed on the theoretical temperatures and densities for the various star models that merit exploration. But we now understand more clearly than before, what sort of things we need to know to clear up the situation.

The Astronomical Conference on the "Internal Constitution of the Stars," held on October 20 and 21, 1939, in New York, under the auspices of The New York Academy of Sciences, considered in turn three specialized subjects.

1. Stellar Opacity.
2. The Density Distribution in Eclipsing Binaries.
3. The Sources of Stellar Energy.

The four papers by Dr. Kopal, Dr. Marshak, Professor Russell, and Dr. Tuominen, included in this article of the Annals of the Academy, represent only a part of the formal contributions from the Conference, since some papers, presented or summarized at the meetings, have been published elsewhere. The informal contributions at this conference of physicists and astronomers were, however, probably the most useful feature. From the off-record consultations and arguments much progress has come. We are in the midst of an interesting and constructive epoch in the study of stellar interiors. The radiative theories have enriched our knowledge of stars and of matter under extreme conditions; and the success of the transmutation hypotheses, in accounting for the sources of stellar energy, have rapidly given us confidence that we are now approaching dependable answers to the question of why the stars shine so constantly and so long.

When inquiries on a subject like stellar structure spread themselves over several branches of science, the importance of a conference, such as that sponsored by The New York Academy of Sciences, becomes important, and almost a necessity. It seems clear that further gatherings of this kind, involving twenty or thirty specialists in more or less informal consultation, should be held in the interests of the advance of astrophysics.

THE DISTRIBUTION OF DENSITY
WITHIN THE STARS

By Henry Norris Russell

Director, Princeton University Observatory, Princeton University, Princeton, New Jersey

To find out anything about the distribution of matter *inside* a star demands the use of an "analytical boring machine"—as Eddington says—and it is only under rather special circumstances that such a machine can get to work. Three different methods, however, have proved available, not merely "in principle" but in practical application. Two of these are of a "classical" nature, and depend essentially upon gravitation; the third involves modern astrophysics. I will speak of this first,—briefly, since it does not really come under our morning's title.

The equilibrium of the gaseous matter inside a star is now pretty well understood. At least, almost all the principal investigators have been led by their researches to agree upon the main principles.

The principal observable characteristics of a star are its mass, its radius, and its luminosity. If we know these, we can find the value of gravity at its surface, the surface temperature, and the rate at which heat flows out per unit area.

We can then work inward from the surface, calculating the density, pressure and temperature in deeper layers, provided that we know two things about the material, the mean molceular weight and the opacity, which determines the rate of rise of temperature with depth. Fortunately for us, both these quantities, under stellar conditions, depend mainly on the percentage of hydrogen which is present. The relative proportions of other kinds of atoms do not matter much (though helium has some influence). Chandrasekhar has developed an effective way of doing this, by which one can figure out with good approximation how deep we would have to go to leave behind us a given fraction (say one-tenth) of the whole mass of the star. The remaining 90 per cent of the mass will then be enclosed in a sphere whose radius is a given fraction (which Chandrasekhar calls ξ^*) of the whole star. For a star of uniform density, ξ^* would be 0.965. Smaller values indicate that the density increases toward the center.

Applying his method to Capella, Chandrasekhar finds that if there is no hydrogen at all inside, $\xi^* = 0.68$, indicating a small central condensation. With 30 per cent of hydrogen it drops to 0.50; with 80 per

cent to 0.32; for still more hydrogen it slowly rises. For Sirius the values ζ* are very slightly smaller; and they are almost the same for the sun, and for most of the stars for which we have reliable data.

Now this hydrogen percentage applies to the outer part of the star; and there is very strong evidence that at the surface itself the outer "atmosphere" is almost all hydrogen. Unless there is an enormous change in the composition with depth, we must conclude that ninety per cent of the mass of these typical stars is contained within a central region between one-half and one-third of the size of the outer surface—that is, that the central condensation is very high. No independent way has, however, yet been found for finding the abundance of hydrogen in the interior of the stars, and hence no definite figures can be given.

Of the two gravitational methods, one depends on the measurement of the shapes of bodies which are distorted by centrifugal or tidal forces. Clairaut, about two hundred years ago, showed that the shape of such a body would depend, not only on the magnitude of this force but also on the degree of condensation toward the center. If φ is the ratio of the distorting force to gravity at the surface, the resulting deviation ε from the spherical form (as a fraction of the radius) is given by the equation

$$= \varphi\left(\frac{1}{2} + K\right). \tag{1}$$

For a body of uniform density, $K = 0.75$. It diminishes with increasing central condensation, and, at the limit (the "Roche model"), when all the mass is concentrated close to the center, $K = 0$.

This principle may be directly applied to most of the planets, since the centrifugal force due to their rotation can be calculated, and the polar flattening ε may be observed. It is thus found that $K = 0.64$ for Mars, 0.47 for the Earth, 0.26 for Jupiter, and 0.12 for Saturn. Mars must therefore be nearly homogeneous, and Saturn far denser at the center than at the surface.

The sun's rotation is unfortunately so slow as to produce no measurable flattening. Many isolated stars are shown to be rotating rapidly by the widening of their spectral lines; but here we have no way to find the shapes of the bodies themselves.

There are many eclipsing binaries—very close pairs, with the stars separated often by less than their own diameters—in which the tidal effect of the attraction of each component upon the other draws them out into prolate forms, with their long axes pointing toward one another.

As the system revolves, the stars are seen alternately broadside on and nearly end on, and a variation of light results, which can usually be well separated from the additional changes due to eclipses.

If it were safe to assume that the star-disks—could we see them sufficiently magnified—would appear uniformly bright all over, our problem would be simple. The observed brightness of each star at any time (barring eclipses) would be simply proportional to the apparent area of the elliptical disk, and it is then very easy to find ε—or, more precisely, a mean value for the two stars, weighted according to their apparent brightness. The theoretical tidal forces are given (to the first approximation) by the simple equations

$$\varphi_1 = 3\frac{m_2}{m_1}\frac{R_1^3}{r^3} \qquad \varphi_2 = 3\frac{m_1}{m_2}\frac{R_2^3}{r^3} \tag{2}$$

where m_1, m_2 are the masses of the stars, R_1, R_2 their radii, and r the distance of their centers. Taking a mean of these, weighted according to the brightness, equation (1) gives a value of K, which again will be a mean for the two components.

Walter, in 1931, concluded in this way that on the hypothesis of uniform star-disks, $\varepsilon = 1.08\varphi$, which gives $K = 0.58$. His observational material was adequate in amount and quality, and it follows that the stars must be nearly homogeneous *provided that this assumption about their apparent disks is correct.*

But there are conclusive reasons for disbelieving this. In the first place, the sun's disk appears much brighter at the center than at the edge. This is to be anticipated theoretically for any gaseous star, in which we can see down to hotter layers when we look straight in. Accurate observations show that it is true for a number of eclipsing variables, and there is no doubt that it must be so generally.

For the total heat-radiation, the brightness at the limb, according to theory, should be about 40 per cent of that at the center. This difference corresponds to a difference of the effective temperatures for the two regions, so that the limb is redder, as well as fainter. Consequently the loss of light at the limb is less than 60 per cent in the infra-red and greater in the ultra-violet, and its exact amount for any given star and given wave-lengths must be found by observation.

It is easy enough to calculate the effect of limb-darkening upon the change in the apparent brightness of a rotating ellipsoidal star; and it is found that it increases this very considerably. The "photometric" ellipticity, previously calculated on the hypothesis of uniform disks, will then be too great, and also the derived value of K.

Theoretical discussion—of the same nature as that which explains the limb-darkening—predicts another important effect. On an ellipsoidal body the surface should shine most brightly in the parts nearest the center, and most faintly at the ends of the long axes. The calculated brightness is proportional to the "local" force of gravity; but this again applies to the total radiation, and the effect in a given wave-length may be greater or less. Moreover, the theoretical discussion involves the assumption that there are no currents causing circulation of matter in the interior of the star, and these, if present, may diminish the effect.

A fairly simple analysis shows that this gravity-effect, too, produces a large increase in the light-variation and in the photometric ellipticity. The latter must be corrected for both gravity-effect and limb-darkening, before the true shape of the stars, and the value of K, can be found.

It should be added that the "reflection" effect, due to the heating of the near side of each star by the radiation of the other, produces a small term which alters the observed photometric ellipticity. But this can be calculated and allowed for (as Walter did). It is no great task to calculate the combined influence of the other three causes. The result is

$$\frac{2C}{\varphi} = \left\{1 + \frac{6x}{15 - 5x}\right\}\left(1 + y + 2K - \frac{1}{2}yK\right) \qquad (3)$$

where C is the observed photometric ellipticity, corrected for reflection effect; x measures the limb-darkening (from 1 at the center to $1 - x$ at the limb) and y is the ratio of the actual gravity-effect for the observed wave-length to its theoretical value. (Note that if $x = y = 0$, this goes over into equation (1) as it should.)

Walter's data give $2C/\varphi = 2.17$.

If we set $x = 0.5$—a rather low value—we find from (3)

$$K = \frac{0.75 - y}{2 - \frac{1}{2}y}.$$

With the theoretical value $y = 1$, K comes out -0.17, an impossible value!

Now there is good evidence from other data that x is not far from the value predicted by theory. For the eclipsing variable, YZ Cassiopeiae, Kron's remarkably accurate observations* give $x = 0.44$.

---

* Pub. A. S. P. **50**: 173. 1938.

But, so far as the writer knows, there is no other method, direct or indirect, for obtaining an observational test of the theory of the gravity-effect. Hence this attractive method of determining the internal constitution of the stars appears to be completely blocked. The worst of it is that even if x and y could be closely estimated from theory, K still appears as the difference of two nearly equal quantities, one derived from observation and the other from the hoped-for theory, so that its determination would in any case be of very low precision.

This is no great tragedy, however. We shall soon see that K can be reliably determined in another way, and the photometric ellipticity then gives information not otherwise accessible about the theoretically important quantity y.

A second gravitational method for determining the central condensation depends upon changes in the orbit of a close binary. It has long been known that a satellite revolving around a spherical planet, in an eccentric orbit, will pursue exactly the same path indefinitely; but if the planet is flattened at the poles, the line of apsides (joining the nearest and remotest points of the orbit) will slowly advance. The rate of this motion depends on the constant K already mentioned. For several of the planets, such changes in the orbits of satellites have been observed, and lead to substantially the same values of K as the ellipticities of figure. For a pair of stars the situation is more complicated. The flattening at the poles, due to rotation, and the elongation toward one another by tidal force, are both operative. The analytical treatment is rather intricate, and has led to some controversy, now satisfactorily cleared up. A very curious difference, however, appears in the solution, according as it is assumed that the stars are solid or fluid. If the latter is the case, and if the fluid is sufficiently mobile to enable the stars to take up, at any moment, substantially the shape corresponding to an instantaneous value of the tidal forces, the line of apsides will advance, as Cowling and Sterne have proved. But if the stars are supposed to be frozen into rigid bodies the results are very different. The two ellipsoids undergo librations—their long axes oscillating about the line joining their centers—and the line of apsides has usually, though not in all cases, a retrograde motion. This conclusion, first reached by Walter, has been confirmed by Sterne. There can be no doubt that the former case represents the actual properties of the stars, which are certainly gaseous throughout.

Cowling's work leads to the equation

$$\frac{n'}{n} = K_1 \left(\frac{R_1}{r}\right)^5 \left(1 + 16\frac{m_2}{m_1}\right) + K_2 \left(\frac{R_2}{r}\right)^5 \left(1 + 16\frac{m_1}{m_2}\right). \qquad (4)$$

Here n is the mean angular motion of the stars in their orbit, and n' that of the line of apsides; the other symbols have the meaning already explained.

To use this equation, we must know the size of each star, as well as of the orbit, and the masses too (or, at least their ratio). Hence we can apply it only in the case of an eclipsing binary. It is hopeless to find the separate values of K for the two stars, but a representative average can be obtained. Sterne has made a more detailed discussion, including several small correction terms, which operate to make the calculated value of K smaller.

For most eclipsing binaries the orbits are practically circular, making them unavailable. But there are some fifteen systems for which the orbits are definitely eccentric, and the eclipse of star A by star B does not come in the middle of the interval between the preceding and following eclipses of star B by star A. So long as the apse does not move this displacement will remain unaltered; but if it does, the intervals between the minima will gradually change. For the systems Y Cygni and GL Carinae, a complete cycle of these changes, corresponding to a whole revolution of the line of apsides, has been observed. The apsidal periods are long—46 and 25 years, while the orbital periods are 3.00 and 2.42 days.

There are several other systems in which the evidence for apsidal motion is definite and periods of from 150 to 600 years are indicated by the observations.

When equation (4) is applied to the seven best observed systems, values of K are found which range from 0.018 to 0.0018. In doubtful cases, the data were so chosen as to lead to the maximum value of K, so that its real mean value is probably somewhat smaller than the mean of the numbers here calculated, namely 0.010.

For several other stars, the existing data are too weak to permit a trustworthy calculation of K, but the apsidal motions are always slow, and it is evident that for these stars, too, K is small. There is not a single case in which it comes out as great as 0.1.

This evidence, which is by far the strongest which exists, indicates that the central condensation of the stars is always great—and fully confirms the conclusion expressed above, that the photometric ellipticity is of value in determining the gravity-effect, but not the condensation.

One may naturally ask, Is there any escape from these conclusions? The existing data are numerous and good enough to leave no loophole so long as equations (3) and (4) are valid. The reasoning on which they are based has been thoroughly checked; but they rest upon one assumption, not yet mentioned: namely, that the periods of rotation and orbital revolution of the stars are the same.

Rotation in a period longer than the revolution seems very improbable—though it may occur for 29 Canis Majoris, which is still incompletely observed. Rotation in a shorter period would increase the polar flattening—thereby causing a more rapid apsidal motion for the same value of K and masses and sizes of the stars,—but might perhaps cause the stars to break loose from the tidal influence, and become spheroidal, instead of ellipsoidal, in figure, thus cutting off about seven-eighths of the effect calculated by Cowling. It is however very doubtful if a mobile mass of gas could fail to yield to the tidal forces— in which case the elongation would be much the same as ever, and also the apsidal motion. This matter deserves further study, and the difficult question what the gravity-effect would be on the surface-brightness of a star rotating in this way has never even been considered. Nevertheless in the writer's judgment, the evidence in favor of a high central condensation is conclusive.

One more word may be said, regarding the form in which these conclusions are often represented. A given value of K tells us that the central condensation is large, or small, as the case may be; but there are a great variety of distributions of internal density which lead to the same value of this constant. Among the vast variety of such possible distributions one sort—the so-called polytropic spheres— has been much discussed, partly because certain reasonable physical assumptions regarding the internal constitution of the stars lead to "models" of this type—and very largely because the properties of these models have been calculated with great accuracy and detail by Emden and others. The various models are characterized by the values of an index n. For $n = 0$, the sphere is of uniform density, for $n = 5$, all the mass is concentrated at the center.

Some of the most important characteristics of these models are recorded in TABLE 1—the ratio $\rho_o/\bar{\rho}$ of the density at the center to the mean density of the whole sphere, the values of the constant K, and of Chandrasekhar's ξ^*.

These models have the advantage that the whole possible range of central condensation can be described by an "index" ranging from 0 to 5. For example, Chandrasekhar's limiting values of ξ^* for

TABLE 1

Polytropic Spheres

n	$\rho_c/\bar{\rho}$	K	ξ^*
0	1.00	0.750	0.965
1	3.29	0.260	0.846
2	11.40	0.074	0.691
3	54.18	0.0144	0.504
4	622 4	0.0013	0.283
5	∞	0 000	0 000

Capella, 0.68 and 0.32 correspond to $n = 2.1$ and $n = 3.8$, while the values found above for K, 0.018 and 0.0018, give $n = 2.9$ and 3.9. Eddington's famous model, derived long ago from over-simplified but reasonable physical assumptions, corresponds to $n = 3$. It is remarkable that after years of rapid progress, the latest data indicate that the stars are actually built in a similar fashion.

SUMMARY

The degree to which the density inside a star increases toward the center can be investigated in three ways.

1. Given the mass luminosity and hydrogen content Chandrasekhar has calculated the thickness of the shell containing the outer tenth of the mass, and found that for the best observed stars the central condensation must be high unless the percentage of hydrogen is very low.

2. Many close binary stars are distorted into ellipsoids by their mutual attraction. The degree of ellipticity depends on the density-concentration, as well as the masses and radii. For many ellipsing pairs there is good photometric evidence of ellipticity. But the relations between this change in brightness and the ellipticity of the stars involve the degree of darkening of the apparent disks toward the limb, and the variation of brightness over their surfaces with change in the local gravity, in such a way that the effects of density concentration are effectively obscured.

3. When such a pair has an eccentric orbit, the line of apsides advances, at a rate depending on the density concentration, as well as the masses and dimensions of the stars. The observed motions are in all cases slow, indicating central densities from 40 to 400 times the mean densities. The central condensation is greater in some stars than others. On the average, it is somewhat greater than Eddington's well-known model.

THE ELLIPTICITY AND REFLEXION EFFECTS IN ECLIPSING BINARY SYSTEMS

By Zdeněk Kopal

From the Harvard College Observatory, Cambridge, Massachusetts

INTRODUCTION

The changes of light displayed by eclipsing binary systems provide us with the possibility of obtaining a vast amount of information concerning their nature. In studying the light curves attention has usually been paid to the conspicuous changes arising from the eclipses. The well known Russell-Shapley method and a number of its subsequent modifications enable one to derive from the light curve, primarily within minima, the geometrical elements of an eclipsing system, namely the radii of the components in terms of the orbital radius and the inclination of the orbital plane to the line of sight. Equally valuable information can, however, be inferred from the additional light changes displayed between eclipses.

Were the variability of an eclipsing binary due to eclipses alone, its luminosity between the two minima should remain constant. In reality, this is approximately true only for very distant systems. As soon as the components move closer, the actual conditions become more complicated, and as a result the maximum brightness ceases to be constant. Two main effects responsible for the variation are known under the terms "ellipticity" and "reflexion" which are self-explanatory. The ellipticity effect, essential for the β Lyrae type of variability, was recognized long ago. The reflection effect, usually far less conspicuous in amount, was discovered and correctly explained by R. S. Dugan[*] in 1908. A simple analytical treatment of both phenomena has been given in the classical papers by Russell and Shapley.[†]

A long series of papers on ellipticity and reflexion in close binary systems has appeared within the past twenty five years. It would be impossible to review them here in full. As regards reflexion, an important advance was made in 1926 when A. S. Eddington[‡] and shortly afterward E. A. Milne[§] re-investigated the theory and physical processes underlying reflexion. The study of the ellipticity effect received

fresh stimulus when H. von Zeipel[*] in a series of remarkable papers showed that for the rotationally or tidally distorted stars the outward flux of radiant energy varies proportionally to the surface gravity. The photometric consequences of this so-called "gravitational effect" upon the light curves of eclipsing binaries were first studied by S. Takeda[†] in an important paper which, unfortunately, long failed to attract attention. The bearing of the gravity effect on some of the problems related to eclipsing variables was further studied by W. Krat,[‡] and, quite recently, the question was independently re-opened by H. N. Russell.[§]

The present paper consists of two parts. In the first section an attempt is made to develop a self-consistent theory of both the ellipticity and reflexion effects, the aim being to provide a basis for constructing theoretical light curves which can be compared with the observations. The second section will be devoted to the application of the theoretical results to practical cases. In developing the theory, we shall assume that the systems under consideration are in a state of equilibrium specified by the following conditions:

1) The eccentricity of the relative orbit of the two components is negligibly small, so that their tidal distortion is practically constant.

2) The components are gaseous configurations, the periods of free oscillation of which are short compared with the period of revolution. The tidal distortion of the surface is then given by the equilibrium theory of tides, and the semi-major axes of both ellipsoids coincide in direction with the radius vector.

3) The axes of rotation of both components are constantly perpendicular to the orbital plane.

There is good observational evidence that the first condition is fulfilled in the vast majority of real eclipsing systems, and equally good theoretical evidence that the second condition will be fulfilled also. Almost nothing, however, is known about the third postulate, and the only reason for introducing it is that it offers considerable simplification in the development of the theory.

It would be interesting to search for possible deviations from the above defined state of equilibrium, as they might reveal a good deal more about the true nature of these systems than the normal cases. However, this problem will be commented upon only briefly at the

---

[*] M. N. **84**: 665, 684. 1924. An independent proof for tidally distorted configurations was also given by Chandrasekhar. M. N. **93**: 571–573. 1933.
[†] Kyoto Mem. A **17**: 197. 1934.
[‡] Z. f. Ap. **9**: 319. 1935.
[§] Ap. J. in press; cf. also P. A. A. S. **9**: 227. 1939.

end, and the question in its complexity must be reserved for future investigation. For the present our interest will be confined to typical systems alone.

THEORETICAL LIGHT CURVES

Ellipticity Effect

The primary cause of the light changes exhibited by close binary systems in addition to the eclipses is the changing area of the visible surface distorted by the mutual gravitational forces of the two components. As we are not considering arbitrary distortions,* but only such as are mutually produced by a pair of gaseous bodies of given mass, structure and relative dimensions, we need to know the geometry of their boundaries in terms of their physical elements. Fortunately, this problem has recently been investigated very fully by S. Chandrasekhar[†] for polytropic configurations. Thus the problem which remains is to investigate the light changes displayed by a pair of such revolving bodies as seen by an observer at a great distance.

GEOMETRY OF THE PROBLEM

The observable luminosity of each component of a binary system can be generally expressed as:

$$L(\phi, \theta) = \int I \cos \alpha \, d\sigma, \tag{1}$$

where I is the flux of radiation per unit area per unit time, and $d\sigma$ the surface element:

$$d\sigma = \xi^2 \sec \beta \sin \theta \, d\theta \, d\phi,$$

the integration being taken over the whole visible hemisphere. The function $\xi(\phi, \theta)$ specifies the actual boundary of the distorted configuration; the angles α and β denote the angle of foreshortening and the angle between the radius vector and surface normal which are given by:

$$\cos \alpha = l.l_0 + m.m_0 + n.n_0, \tag{2.1}$$

$$\cos \beta = \lambda.l + \mu.m + \nu.n, \tag{2.2}$$

where

$$\lambda, \mu, \nu$$
$$l, m, n$$
$$l_0, m_0, n_0$$

---

* From a purely geometrical point of view, the light changes displayed by rotating ellipsoids of given form were recently investigated by K. Schütte. A. N. **267**: 369. 1939.

† M N. **93**: 390, 449, 462, 539. 1933.

are direction cosines of the radius vector, surface normal, and line of sight respectively.

The values of λ, μ, ν can be put down at once:

$$\lambda = \cos \phi \sin \theta,$$
$$\mu = \sin \phi \sin \theta,$$
$$\nu = \cos \theta.$$

The values of l, m, n are specified by the form of the free surface of the distorted body. This is obviously an equipotential defined as:

$$W = \text{constant}, \tag{3}$$

where W is the combined potential of forces acting on the surface, *i. e.*, the potential arising from the mass of the distorted component, the centrifugal potential (if it exists), and the tide-generating potential due to the disturbing secondary. The direction cosines of the surface normal at any point then are:

$$l = \frac{1}{N}\frac{\partial W}{\partial x}, \tag{4.1}$$

$$m = \frac{1}{N}\frac{\partial W}{\partial y}, \tag{4.2}$$

$$n = \frac{1}{N}\frac{\partial W}{\partial z}, \tag{4.3}$$

where

$$N^2 = \left(\frac{dW}{dn}\right)^2 = \left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2 + \left(\frac{\partial W}{\partial z}\right)^2.$$

Explicitly,

$$l = \lambda\{1 - 2u_2 + 2u_2\lambda^2 - 2v\,\nu^2 + \ldots\ldots\}, \tag{4.11}$$

$$m = \mu\{1 \qquad + 2u_2\lambda^2 - 2v.\nu^2 + \ldots\ldots\}, \tag{4.21}$$

$$n = \nu\{1 + 2v \quad + 2u_2\lambda^2 - 2v.\nu^2 + \ldots\ldots\}, \tag{4.31}$$

where we have abbreviated

$$u_j = \frac{3}{2}\Delta_j \frac{m'}{m}\, r^{j+1}, \tag{5}$$

$$v = \frac{1}{2}\Delta_2\left(1 + \frac{m'}{m}\right)r^3.$$

Here m/m' denotes the mass-ratio of the disturbing to the distorted body, r the radius of a sphere having the same volume as the distorted ellipsoid and expressed in terms of the orbital radius taken as unity, and the Δ_j's are factors depending on the internal density distribution introduced by Chandrasekhar,[*] and, for polytropic configurations, defined as:

$$\Delta_j = \frac{(2j + 1)\,\Psi^{\circ}_j(\xi_1)}{(j + 1)\,\Psi^{\circ}(\xi_1) + \xi_1\Psi'(\xi_1)}. \tag{6}$$

Here ξ_1 is the first zero solution of the Emden equation for polytropic index n; the $\Psi_j(\xi_1)$'s are surface values of associated Emden functions of the j'th order (Chandrasekhar*) at the boundary of initially undistorted configurations, and the Ψ_j''s are their derivatives. The numerical values of Δ_j for different polytropic indices from 0 to 5 are given in TABLE 1.

The values of the remaining direction cosines l_0, m_0, n_0 depend on the observer's position. If ψ denotes the phase angle and i the inclination of the orbital plane to one perpendicular to the line of sight, these values reduce to:

$$l_0 = \cos\psi\sin i,$$
$$m_0 = \sin\psi\sin i,$$
$$n_0 = \cos i.$$

TABLE 1

n	ρ_c/ρ_m	Δ_2	Δ_3	Δ_j
0	1.0000	2.50000	1.75000	1.50000
1	3.2899	1.51982	1.21291	1.12048
1.5	5.9907	1.2892	1.1079	1.0562
2	11.4025	1.1482	1.0488	1.0231
3	54.1825	1.0289	1.00736	1.00281
4	622.408	1.00267	1.00047	1.00014
5	∞	1.00000	1.00000	1.00000

The fundamental equation (1) then takes the form:

$$L(\psi, i) = \int_0^\pi \int_{\psi - \frac{\pi}{2}}^{\psi + \frac{\pi}{2}} \xi^2\, I \cos\alpha \sec\beta \sin\theta\, d\theta d\phi. \tag{7}$$

GRAVITATION EFFECT

The expression for I remains to be derived. If the brightness over the surface is constant, the integration of (1) can be performed at once and the result is well known[†]. The reality, however, is likely to be more complex. First, the researches by von Zeipel mentioned above have led to the conclusion that the outward flux of radiant energy

---

* M. N. 93: 453. 1933.
† Russell. Ap. J. 36: 385. 1912.

per unit area at any point across the boundary of a rotationally or tidally distorted configuration should vary proportionally as the surface gravity. Thus:

$$I = c\frac{dW}{dn},\tag{8}$$

which would evidently be true for the total radiation only. If we assume that our configuration radiates like a black body, and if $E(\lambda, T)$ denotes its radiation for effective temperature T and wave length λ, while $E(T)$ is its integrated light, equation (8) specifies the variations of temperature on the distorted surface, *i. e.*:

$$E(T) \equiv a.T^4 = c'\frac{dW}{dn},\tag{9}$$

and the intensity of radiation observed in any particular wave length λ becomes:

$$\frac{I}{I_0} = \frac{E(\lambda, T')}{E(\lambda, T)} = \frac{e^{\frac{c_2}{\lambda T}} - 1}{e^{\frac{c_2}{\lambda T'}} - 1},\tag{10}$$

where $c_2 = 1.432$ cm. grad, T is the mean effective temperature of the whole disc, and T' is the "local" temperature at any point on the distorted boundary. In practical cases we may assume that T' will never be radically different from T, and may expand (10) in terms of $(T/T' - 1)$. For this we need to know equation (9) in its explicit form. From Chandrasekhar's investigations[*] we may take:

$$\frac{dW}{dn} = \frac{dW_0}{dn}\left\{1 + \cdots - \frac{2}{3}\sum_{j=2}^{4}\left(\frac{2j + 1}{\Delta_j} + 1 - j\right)u_jP_j(\lambda)\right\},\tag{11}$$

where dW_0/dn is the (constant) value of gravity at the surface of an initially undistorted sphere, and the $P_j(\lambda)$'s are surface harmonics. This expression is correct as far as terms of the order of r^5 are concerned. As, in view of (7), only terms depending on λ will contribute to the phase effect, purely radial terms or those involving ν are omitted from (11) as immaterial.

Inserting (11) in (9) and expanding (10) in terms of u, we find:

$$I = I_0\left\{1 + \cdots - \frac{2}{3}\tau\sum_{j=2}^{4}\left(\frac{2j + 1}{\Delta_j} + 1 - j\right)u_jP_j(\lambda)\right\},\tag{12}$$

where

---

[*] M. N. **93:** 570. 1933.

$$\tau = \frac{E(T)}{E(\lambda, T)} \cdot \frac{dE(\lambda, T)}{dE(T)} = \frac{1}{4} \frac{c_2}{\lambda T\left(1 - e^{-\frac{c_2}{\lambda T}}\right)}. \tag{13}$$

The numerical values of the quantity 4τ, needed frequently, are given for different λ and T in TABLE 2. If we compare equations (11) and (12), we note that if (to our order of accuracy) $\tau = 1$, $i.\,e.$, if

$$\lambda.T = 2.738 \ldots, \tag{14}$$

equation (12) expresses the intensity distribution of the total radiation.

LIMB DARKENING

The quantity I_0 from (12) could be regarded as constant, were it not for the effect of limb darkening. It is well known that this is caused by the fact that radiation emerging from the stellar atmosphere in different directions originates on the average at a different optical depth, and therefore corresponds to a different temperature if we pass from center to limb. The variation of temperature is given as:[*]

$$T_\alpha{}^4 = T_0{}^4\left(1 + \frac{3}{2}\frac{\bar{\kappa}}{\kappa_\lambda}\cos\alpha\right), \tag{15}$$

where T_0 is the mean surface temperature of the star, related to its effective temperature by:[†]

$$\left(\frac{T}{T_0}\right)^4 = \frac{4}{\sqrt{3}}, \tag{16}$$

while T_α is the surface temperature at a point where the line of sight makes an angle α with the surface normal. As is evident, α is identical with the angle of foreshortening. Further, κ_λ denotes the mass-absorption coefficient for the wave length λ, and $\bar{\kappa}$ is the continuous absorption coefficient for the total radiation (Rosseland mean).

In consequence, as in the preceding paragraph,

$$\frac{I_0}{H} = \frac{E(\lambda, T_0)}{E(\lambda, T_\alpha)} = 1 + \frac{3}{2}\tau_0\frac{\bar{\kappa}}{\kappa_\lambda}\cos\alpha, \tag{17}$$

where τ_0 is identical with τ as defined by equation (13) except that the effective temperature in (13) is to be replaced by the surface tempera-ture. If, further, we put

---

[*] *E.g.* Eddington, "Internal Constitution of the Stars." p. 320. 1930.

[†] Bronstein, Phys. Zs. 59: 144. 1929; also Hopf, Cambridge Math. Tracts No. 31. 1934.

TABLE 2

λ(in Å) / $10^{-?}\cdot T$	3750	4000	4250	4500	4750	5000	5250	5500	6000	7000	8000	10000
4	9.547	8.951	8.423	7.956	7.539	7.166	6.827	6.519	5.983	5.145	4.527	3.083
5	7.632	7.166	6.747	6.374	6.044	5.747	5.479	5.236	4.813	4.160	3.083	3.037
6	6.355	5.982	5.636	5.331	5.236	4.814	4.595	4.396	4.054	3.526	3.142	3.269
7	5.479	5.330	4.852	4.595	4.366	4.160	3.978	3.812	3.526	3.088	2.853	2.350
8	4.814	4.527	4.275	4.054	3.858	3.683	3.526	3.386	3.142	2.772	2.505	2.149
9	4.305	4.054	3.835	3.642	3.483	3.320	3.184	3.062	2.853	2.535	2.304	1.998
10	3.902	3.683	3.467	3.320	3.170	3.037	2.919	2.811	2.629	2.350	2.149	1.881
12	3.320	3.142	2.987	2.852	2.733	2.629	2.534	2.450	2.304	2.084	1.925	1.712
14	2.919	2.853	2.645	2.534	2.436	2.350	2.273	2.202	2.084	1.902	1.772	1.597
16	2.629	2.505	2.399	2.304	2.222	2.149	2.084	2.025	1.925	1.772	1.662	1.513
18	2.410	2.304	2.212	2.132	2.107	1.998	1.942	1.891	1.805	1.674	1.579	1.450
20	2.239	2.149	2.086	1.997	1.936	1.881	1.832	1.788	1.712	1.597	1.513	1.400
24	1.999	1.925	1.861	1.805	1.756	1.712	1.673	1.639	1.576	1.486	1.419	1.328
28	1.833	1.772	1.719	1.674	1.683	1.597	1.565	1.536	1.486	1.410	1.351	1.277
35	1.642	1.597	1.461	1.522	1.491	1.464	1.440	1.418	1.379	1.320	1.277	1.218
50	1.429	1.400	1.375	1.351	1.331	1.314	1.298	1.283	1.257	1.218	1.190	1.150

$$u = \cfrac{1}{1 + \cfrac{2}{3} \cdot \cfrac{1}{\tau_0} \cdot \cfrac{\kappa_\lambda}{\kappa}} \tag{18}$$

equation (17) takes the usual form:

$$I_0 = H(1 - u + u \cos \alpha), \tag{19}$$

The constant H can then be determined from the equation:

$$4L = \int I d\sigma, \tag{20}$$

where $4L$ represents the total luminosity, and the integration is to be extended over the whole surface.

PHASE EFFECT

The final step is to evaluate (7). In order to do so, we need an explicit expression for ξ, which likewise can be taken from Chandrasekhar.[*] Omitting again radial and centrifugal terms we find:

$$\xi = \xi_1 \left\{ 1 + \ldots \ldots + \frac{2}{3} \sum_{j=2}^{4} u_j P_j(\lambda) \right\}. \tag{21}$$

If we insert this and equations (13) and (19) in (7), the final integration presents no difficulties. The integral can be expanded in powers of l_0:

$$L\left(\psi, i\right) = L\left(\frac{\pi}{2}, i\right) \left\{ 1 + \frac{m'}{m} \sum_{j=1}^{\infty} c_j \cos^j \psi \sin^j i \right\}, \tag{22.0}$$

and the coefficients take the forms:

$$c_1 = + \frac{5u}{2(3 - u)} \left[\Delta_3 + \frac{\tau}{4}(7 - 2\Delta_3) \right] r^4 + \ldots \ldots \tag{22.1}$$

$$c_2 = - \frac{3}{10} \frac{15 + u}{3 - u} \left[\Delta_2 + \frac{\tau}{4}(5 - \Delta_2) \right] r^3 - \frac{15}{8} \frac{1 - u}{3 - u} \left[\Delta_4 + \right.$$
$$\left. \frac{\tau}{4}(9 - 3\Delta_4) \right] r^5 + \ldots \tag{22.2}$$

$$c_3 = - \frac{25}{6} \frac{u}{3 - u} \left[\Delta_3 + \frac{\tau}{4}(7 - 2\Delta_3) \right] r^4 + \ldots \ldots \tag{22.3}$$

$$c_4 = + \frac{35}{16} \frac{1 - u}{3 - u} \left[\Delta_4 + \frac{\tau}{4}(9 - 3\Delta_4) \right] r^5 + \ldots \ldots \tag{22.4}$$

[*] M. N. 93: 467. 1933.

The maximum luminosity at quadrature, $L(\frac{\pi}{2}, i)$, is conventionally taken as unity.

One can easily verify that as long as there is no limb darkening, the expansion in (22) contains only even powers of l_0 (the gravity effect being symmetrical with respect to the radius vector). The appearance of limb darkening (symmetrical to the line of sight) introduces also odd powers of l_c, with co-efficients of the order of r^4 and higher.

COMPARISON WITH PREVIOUS RESULTS

If in the above formulae for the coefficients we put arbitrarily $u = 0, \tau = 0$, we obtain the expression valid for uniform discs, and equation (22) reduces to:

$$L(\psi, i) = L\left(\frac{\pi}{2}, i\right)\left\{1 - \frac{\varepsilon^2}{2}\cos^2\psi\sin^2 i + \ldots\ldots\right\}, \quad (23.1)$$

where ε^2, the dynamical ellipticity, is:

$$\varepsilon^2 = 3\frac{m'}{m}r^3\left[\Delta_2 + \frac{5}{12}\Delta_4 r^2 + \ldots\right]. \quad (23.2)$$

Equation (23.1) in this form was first derived by Russell.[*] It can, however, never be of astrophysical importance. For high temperature stars, u may well tend to zero, but τ can never become less than one-fourth. Consequently the "optical ellipticity," *i. e.*, the photometrically observed value of the coefficient c_2, will in any case be greater than the dynamical one.

Putting $u = \frac{3}{5}$ and $\tau = 1$, we get formulae relevant to the total radiation. If, also, we limit our attention to the centrally condensed configurations ($\Delta_j \sim 1$), we find that, omitting the limb darkening, allowance for the gravity effect makes the optical ellipticity just twice as large as the dynamical ellipticity, *i. e.*, within the scheme of our approximation the observed luminosity will change as

$$L(\psi, i) = L\left(\frac{\pi}{2}, i\right)\left\{1 - \varepsilon^2\cos^2\psi\sin^2 i + \ldots\right\}. \quad (23.3)$$

This important result was first pointed out by Takeda.[†] If we allow also for the darkening, the magnifying factor increases from 2 to 2.6.

[*] Ap. J. **36**: 385. 1912. The occurrence of the term of the order of r^4 was first noticed by Klauder (A. N. **255**: 1. 1935), on the basis of Chandrasekhar's results.
[†] Kyoto Mem. A **12**: 212. 1934.

In reality, however, τ may be either greater or smaller than one. A glance at (14) shows that if $c_2/\lambda T$ is small, τ approaches one-fourth; if $c_2/\lambda T$ is large, τ varies nearly proportionally to it. Therefore, given a dynamical ellipticity, the optical ellipticity will appear the greater, the lower the surface temperature, or the shorter the effective wave length. It becomes equal to the optical ellipticity appropriate for a total radiation, if λ and T fulfill equation (15).

The bearing of the gravitational effect upon the optical ellipticities has recently been investigated also by H. N. Russell.[*] His treatment is relevant only to ellipsoidal configurations (*i. e.*, it includes terms of the order of r^3), and to this order of accuracy his results are perfectly consistent with ours.[†] Verifying that Russell's k is equal to

$$k = \tfrac{1}{2}(\Delta_2 - 1),$$

we find that his adjustable constant y, for black-body radiation, becomes

$$y = \tau. \tag{23.4}$$

Equation (22) derived in the present paper is complete as far as distortions of the order of r^5 are concerned, and applicable to observations of stars of any temperature in any effective wave length. The limit of its applicability in astrophysical problems will essentially depend only on how closely the actual stellar radiation comes to that of a black body.

Reflexion Effect

It is inevitable in close binary systems that a part of the radiation of either component will fall on the surface of its companion, where it will be absorbed and re-emitted. The amount of additional radiation thus contributed to the total observed light of the system will clearly depend on both the geometrical and physical properties of such a system, and will vary with the phase. In order to be able to evaluate it, we must first know the albedo of the "reflecting"[‡] surface. A necessary consequence of the assumption of radiative equilibrium is that the incident radiation must be wholly re-emitted, *i. e.*, that the heat-albedo must be unity.[§] A simple geometrical consideration shows, further, that the amount of reflected light will, to the first

---

[*] In press.

[†] The writer is greatly indebted to Professor Russell for kindly permitting him to see his paper in manuscript.

[‡] This term is here used in the technical sense only, for the true mechanism of the "reflexion" effect is re-emission.

[§] For the general argument, Eddington, M. N. 86: 320; or Milne, M. N. 87: 44. 1926.

approximation, be proportional to the exposed area, *i. e.*, to a quantity of the order of r^2. The effect of distortions will thus appear in terms starting from r^5. There would be no difficulty in taking them into account; the procedure developed in a previous section would enable one to include terms up to the order of r^7. But there is another reason which would render such refinements rather useless, namely, the scattering of incident radiation. If the latter propagates in a direction normal to the surface of the emitting star, which we shall for the sake of brevity call the primary, its radius (as far as the primary can be regarded as spherical) is irrelevant. If, however, the light is scattered in the primary's atmosphere and the radiation from any point propagates in any direction, the radius of the primary does enter, as well as its degree of darkening. Unfortunately, our present knowledge of the proportion of scattered to total light in stellar atmospheres is very incomplete, and as soon as the effect of scattering appears a definite uncertainty enters—which makes it necessary not to go beyond terms of the order of r^4. In order to avoid it, therefore, a theory of the reflexion effect will be developed, including terms of the order of r^4—valid as far as both components can be regarded as spherical, and irrespective of scattering.

GEOMETRY OF THE PROBLEM

The amount of light re-radiated by either component can generally be expressed as:

$$L_j^*(\psi, i) = \int I^* \cos \alpha \, d\sigma, \tag{24}$$

where I^* is the intensity of the reflected radiation, and the integration is to be extended over the whole visible crescent. An attempt at a general solution of (24) for any value of ψ and i leads to very troublesome algebra. An essential condition for the occurrence of eclipses in a binary system is, however, that the inclination of the orbital plane to the line of sight be small. If $i = \frac{1}{2}\pi$, the geometry becomes largely simplified and (24) assumes the form

$$L_j^*\left(\psi, \frac{\pi}{2}\right) = \int_0^\pi \int_{\pi/2-\psi}^{\pi/2} \xi_j^2 I^* \cos \alpha \sec \beta \sin \theta \, d\theta d\phi, \tag{24.1}$$

the evaluation of which presents no difficulties.

The intensity of reflected radiation I^* is, according to Milne,[†] given by

$$I^* = \frac{H'}{\pi}\left[\frac{(\cos\alpha + \tfrac{1}{2})\,(\cos\gamma + \tfrac{1}{2})}{\cos\alpha + \cos\gamma}\right]\cos\gamma. \qquad (25)$$

Here

$$\pi H' = L_i\{1 + r_i P_1(\lambda) + r_i^2 P_2(\lambda) + \ldots\}^2, \qquad (26)$$

L_i being the luminosity of the illuminating star, and γ denoting the angle between the direction of incident radiation and the surface normal which can be expressed as

$$\cos\gamma = \lambda + (\lambda^2 - 1)\,r_j + \frac{3}{2}\lambda\,(\lambda^2 - 1)\,r_j^2 + \ldots. \qquad (27)$$

As is evident from equation (26), the angles α and γ are practically identical for fairly separated systems; for closer systems $\gamma > \alpha$, but the relationship between the two angles is such that the expression in brackets in (25) is very nearly equal to unity in most of the range that concerns us. Therefore in what follows we put the term in brackets equal to unity; *i. e.*, we assume that within the scheme of our approximation the true intensity distribution can be replaced by Lambert's law.[‡]

The integral of (24.1) can now be solved at once:

$$L_j^*\left(\psi, \frac{\pi}{2}\right) = \frac{2}{3\pi}\,L_i r_j^2\{\sin\psi - \psi\cos\psi\} +$$

$$+ \frac{3}{8}\,L_i r_j^3\left\{\cos^2\psi - \frac{2}{3}\cos\psi - \frac{1}{3}\right\} - \qquad (28)$$

$$- \frac{1}{\pi}\,L_i r_j^4\{\sin^3\psi\} + \ldots.$$

If we retain only the first term of the order of r^2 (*i. e.*, if we consider the beams of incident and reflected light as parallel), the result is one derived long ago in connection with planetary phases.[§] If, however, the two bodies approach closer (as in eclipsing systems), this approximation becomes obviously inadequate and higher terms have to be taken into account.

---

† M N. 87: 43. 1926.

‡ This approximation obviously fails at the edge where, according to the cosine law, the intensity should vanish, whereas in reality it remains finite (*i. e.*, Lambert's law corresponds to the degree of darkening $u = 1$, whereas the more correct law (25) yields $u = 0.9$). But near the edge the intensity of reflected radiation becomes so small that its contribution to the total reflected light can be neglected.

§ Schoenberg, Theoretische Photometrie, Handb. d. Astrophysik, First Half, Part 2. 1: 64.

PHASE EFFECT

Let us expand equation (28) in a Fourier series in terms of $\cos n\psi$ (n being an integer). If we remember that $L^*(\psi, \frac{\pi}{2})$ is a single-valued function of ψ in the interval $0 > \psi > \pi$, and if we verify that for $n = 1$ the only non-vanishing term is

$$\int_0^\pi \psi \cos^2 \psi d\psi = \frac{\pi^2}{4},$$

while for $n > 1$ we have

$$\int_0^\pi (\sin \psi - \psi \cos \psi) \cos n\psi d\psi = \begin{cases} \dfrac{4}{9.15 \ldots (3n+3)}, & \text{if } n \text{ is even;} \\ 0, & \text{odd;} \end{cases}$$

and

$$\int_0^\pi \sin^3 \psi \cos 2\psi d\psi = -\frac{4}{5\pi},$$

while, for $n > 2$,

$$\int_0^\pi \sin^3 \psi \cos n\psi d\psi = \begin{cases} \dfrac{4}{5.7 \ldots (n+3)} & \text{even,} \\ 0 & \text{odd,} \end{cases} \text{ if } n \text{ is}$$

equation (28) takes the form

$$L_i^*\left(\psi, \frac{\pi}{2}\right) = -\frac{1}{3}L_i\left\{r_j^2 \qquad + \frac{3}{4}r_j^3 + \ldots \qquad\qquad\right\} \cos \psi +$$

$$+ \frac{1}{3}L_i\left\{\frac{16}{9\pi^2}r_j^2 \quad + \frac{9}{16}r_j^3 + \frac{24}{5\pi^2}r_j^4 + \ldots\right\} \cos 2\psi + \quad (29)$$

$$+ \frac{1}{3}L_i\left\{\frac{16}{135\pi^2}r_j^2 + \qquad - \frac{24}{35\pi^2}r_j^4 + \ldots\right\} \cos 4\psi + \ldots$$

Rewriting it in terms of powers of $\cos \psi$ we obtain

$$L_i^*\left(\psi, \frac{\pi}{2}\right) = -\frac{1}{3}L_i\left\{r_j^2 + \frac{3}{4}r_j^3 + \ldots\right\} \cos \psi + \qquad (30)$$

$$+ \frac{1}{3}L_i\left\{\frac{32}{\pi^2}\sum_{n=2}^\infty \frac{(-1)^n (n-1)^2}{9.15 \ldots (6n-3)}r_j^2 + \frac{9}{8}r_j^3 + \right.$$

$$\left. \frac{48}{\pi^2}\left(\frac{1}{5} + \sum_{n=3}^\infty \frac{(-1)^{n-1}(n-1)^2}{5.7 \ldots (2n+1)}\right)r_j^4 + \ldots\right\} \cos^2 \psi$$

$$+ \frac{1}{3}L_i \left\{ \frac{8.16}{9.15\pi^2} \left(\frac{15}{21} + \sum_{n=2}^{\infty} \frac{(-1)^n (2n+4).5.6\ldots(n+3)}{n! \, 21.27 \ldots (6n+9)} \right) r_j^2 - \right.$$

$$- \frac{8.24}{35\pi^2} \left(\frac{1}{3} + \sum_{n=2}^{\infty} \frac{(-1)^n (2n+4) \, 5.6 \ldots (n+3)}{n! \, 9.11.13 \ldots (2n+5)} \right) r_j^4 + \ldots \right\} \cos^4 \psi + \ldots$$

When we evaluate the numerical coefficients, equation (30) finally takes the form:

$$L_j^*\left(\psi, \frac{\pi}{2}\right) = - L_i \{0{\cdot}3333 \, r_j^2 + 0{\cdot}2500 \, r_j^3 + \ldots \qquad \} \cos \psi +$$

$$\text{(30.1)} \qquad + L_i \{0{\cdot}0925 \, r_j^2 + 0{\cdot}3750 \, r_j^3 + 0{\cdot}4698 \, r_j^4 + \ldots\} \cos^2 \psi +$$

$$+ L_i \{0{\cdot}0239 \, r_j^2 + \qquad\qquad - 0{\cdot}0928 \, r_j^4 + \ldots\} \cos^4 \psi +$$

$$\cdots$$

If terms arising from r^3 and r^4 are neglected, one can easily verify that equation (29) checks with Eddington.[†] A more rigorous law, valid if proper darkening of the reflected radiation is taken into account, was derived by Milne[‡] to the first order of approximation and is substantially the same.

So far $i = \frac{1}{2}\pi$ has been assumed throughout. If its deviation from 90° is small, one can infer from the general form of our equations that, approximately,

$$L^*(\psi, i) = L^*\left(\psi, \frac{\pi}{2}\right) \sin i.$$

[†] M. N. 86: 320. 1926. For the amount of light reflected in full phase Eddington derived the more exact expression

$$L_j^* \left(\pi, \frac{\pi}{2} \right) = \frac{2}{3} L_i \left\{ \sin^2 \phi_0 + (2 + \cos^3 \phi_0 - 3 \cos \phi_0)/\sin \phi_0 \right\} \cdot$$

where $\phi_0 = \sin^{-1} r_j$. Expanding the expression in brackets in powers of r_j, we obtain

$$L_j^* \left(\pi, \frac{\pi}{2} \right) = \frac{2}{3} L_i \left(r_j^2 + \frac{3}{4} r_j^3 + \frac{1}{4} r_j^5 + \frac{9}{64} r_j^7 + \ldots \right),$$

which, to terms of the order of r^4, agrees with the present result. Eddington's higher terms would be true only if both bodies were rigid and if there were no scattering of light.

[‡] Milne, using equation (25) instead of Lambert's law, derived the reflexion law

$$L_j^* \left(\psi, \frac{\pi}{2} \right) = L_i r_j^2 \left\{ \frac{\sin \psi - \psi \cos \psi}{3\pi} + \frac{1 - 5 \cos \psi}{16} + \frac{(1 + \cos \psi)(1 + 3 \cos \psi)}{32 \sin (\psi/2)} \log \frac{1 + \sin (\psi/2)}{\cos (\psi/2)} \right\}.$$

which, expanded in a Fourier series, becomes

$$L_j^* \left(\frac{\pi}{2}, \psi \right) = -\frac{1}{3} L_i \left\{ 1.041 \cos \psi - 0.327 \cos 2\psi + 0.015 \cos 3\psi - 0.003 \cos 4\psi + \ldots \right\} r_j^2$$

For i deviating largely from 90° the exact amount of reflected light becomes uncertain, but it is always smaller than $L^*\left(\psi, \dfrac{\pi}{2}\right)$.

REFLEXION OBSERVED IN A GIVEN WAVE LENGTH

The formulae developed thus far are based on the assumption that the albedo of the "reflecting" star is unity. This is obviously true only for the integrated radiation, and our formulae as they stand would therefore be generally applicable only to radiometric light curves. If observations in a more or less narrow spectral range are available, however, and the two components are of different spectral types, the incident radiation is re-radiated at a different temperature, and the results obtained so far need to be corrected because the efficiency of the re-radiated light at any particular wave length will generally be different from that of the incident radiation.[†]

The appropriate expression for light curves observed in the effective wave length λ is easily found to be

$$L_{j\lambda}^*(\psi, i) = \frac{E(T_i)}{E(T_j)} \cdot \frac{E(\lambda, T_j)}{E(\lambda, T_i)} \cdot L_{j\,\text{bol}}^*(\psi, i). \tag{31}$$

The ratio $E(\lambda, T_j)/E(\lambda, T_i)$ is a matter of observation and can be determined for eclipsing systems from the observed loss of light in both minima. For specifying the ratio between the bolometric brightnesses, recourse must be had to a definite law of radiation. The black-body law which can safely be taken as a good approximation yields

$$\frac{E(T_i)}{E(T_j)} = \left(\frac{T_i}{T_j}\right)^4, \tag{32}$$

and

$$\frac{E(\lambda, T_j)}{E(\lambda, T_i)} = \frac{e^{\frac{c_2}{\lambda T_i}} - 1}{e^{\frac{c_2}{\lambda T_j}} - 1}. \tag{33}$$

Both ratios turn out to be dependent on the temperatures of both components, and the latter to be dependent also on the effective wave length in which the light curve was observed.

It should be pointed out that the ratios of temperatures in (32) or (33) refer to illuminated hemispheres and are likely not to be identical with the proper temperatures of the two stars. The incident radiation produces a definite heating effect tending to increase, in propor-

---

† **Kopal**, Ap. J. **89**: 323. 1939.

tion to its intensity, the temperatures of the illuminated hemispheres, *i. e.*, to equalize them. The theory of the heating effect has been developed by Milne[†] who found that in the presence of incident radiation the true surface temperature of the "reflecting" star T_r is, very approximately,[‡]

$$T_r{}^4 = T^4(1 + 2\rho), \tag{34}$$

where

$$\rho = \frac{\text{reflected light}}{\text{ordinary light}}. \tag{35}$$

The term "reflected light" refers here to the total radiation; hence

$$\rho_i = \frac{2}{3}\frac{L_j}{L_i}\left\{r_i{}^2 + \frac{3}{4}r_i{}^3 + \dots\right\}. \tag{35.1}$$

But, in view of Stefan's law,

$$\frac{L_i}{L_j} = \left(\frac{r_i}{r_j}\right)^2\left(\frac{T_i}{T_j}\right)^4.$$

Remembering that the ratios of effective and surface temperatures are identical, we may write

$$\left(\frac{T_i}{T_j}\right)_r = \frac{T_i}{T_j}\left\{\frac{1 + a_i\dfrac{T_i}{T_j}}{1 + a_j\dfrac{T_j}{T_i}}\right\}^{\frac14}, \tag{36}$$

where

$$a_i = \frac{4}{3}r_i{}^2\left(1 + \frac{3}{4}r_j + \dots\right). \tag{36.1}$$

With the aid of equation (36), the ratios of the temperatures of the illuminated sides can be reduced to the proper temperatures of the stars and vice versa.

COMPARISON WITH OBSERVATIONS

In applying the formal results of the above analysis to reality, one obvious circumstance must be taken into account. In all eclipsing systems the observed light curves represent the superposed effects of two (or sometimes more) components. What we actually observe is not the reflexion of the primary's light on the secondary only, but the

---

[†] M. N. **87: 43. 1926.**
[‡] *i. e.*, if we neglect the darkening of the reflected radiation. Taking it into account, Milne found that the coefficient of ρ becomes, more accurately, 1.95 instead of two.

combined inter-reflexion of light between the two components. The convex form of the light curves between minima does not generally reveal the form of any individual component, but specifies the so-called mean ellipticity of a system. As is evident from the theory and has long been recognized,[†] the dynamical ellipticities of the components in a binary system are generally different, the observable mean ellipticity being the mean value of the proper ellipticities of the two components, weighted according to their luminosities.[‡] Fortunately, the weights with which the individual ellipticities enter into the mean are known, and hence the relation between the individual and mean ellipticities can be established. Forming the means and combining equations (22) and (30) we can now, within the limits outlined in the introduction, predict the theoretical form of the light curve of a given eclipsing system between the minima. As will be seen, the comparison of the theory with observation will enable us— *inter alia*—to obtain information about certain otherwise hardly accessible quantities of high astrophysical interest, *viz.*, the degrees of limb darkening, or the degrees of stellar density condensation.

Harmonic analysis of the light curves between minima proceeds usually by the assumption that the light changes are expansible in the form

$$L(\psi) = 1 - s \cos \psi - z \cos^2 \psi + \ldots. \qquad (37)$$

and n normal points of the light curve between minima provide us with the same number of equations of the above type, a least squares solution of which yields the values of the constants s and z with their probable errors. Coefficients of higher terms in the above expansion, if derived, are usually too small to be significant. The first two coefficients s and z are conventionally called "reflexion" and "ellipticity" constants, according to the nature of their origin. Although we have seen in the preceding part that in the higher order terms both effects superpose, we shall retain these notations for practical reasons and shall discuss the interpretation of each coefficient in turn.

Reflexion Effect

The equation (30) governs the variation in reflected light with phase. It contains one term varying as $\cos \psi$. Accordingly, the primary's light reflected on the secondary gives rise to the term

---

[†] Myers, G. W. Ap. J. **7**: 8. 1898.

[‡] The question how closely the light changes of a system consisting of two different ellipsoids, revolving about their common center of gravity, can be approximated by the light changes of a system composed of two similar ellipsoids has recently been studied in great detail by H. N. Russell (in press).

$$- \frac{L_1}{3} \frac{J_2}{J_1} \left(\frac{T_1}{T_2}\right)^4 \left(r_2{}^2 + \frac{3}{4} r_2{}^3 + \ldots\right) \sin i,$$

and the amount of the secondary's light reflected on the primary is obtained by interchanging the indices and altering the phase by π. The cosine changes its sign and, in consequence, the effects arising from either component tend to cancel. What we can actually observe is only the difference in the light reflected by the two components.

Further, in studying the ellipticity effects we found that the axial rotation of a distorted ellipsoid darkened to the limb gives rise also to another term varying as $\cos \psi$, which we denoted c_1 (equation 22.1). Eliminating J_1/J_2 by means of the relation $J_1/J_2 = (r_2/r_1)^2$. L_1/L_2, we find that

$$s = (s_r - s_e) \sin i, \tag{38}$$

where

$$s_r = \frac{1}{3} \left\{ L_2 r_1{}^2 \left(1 + \frac{3}{4} r_2 + \ldots\right) \left(\frac{T_1}{T_2}\right)^4 - L_1 r_2{}^2 \left(1 + \frac{3}{4} r_1 + \ldots\right) \left(\frac{T_2}{T_1}\right)^4 \right\}, \tag{39}$$

$$s_e = \frac{5}{2} \left\{ \frac{L_1 u_1}{3 - u_1} \left[\Delta_3 + \frac{\tau_1}{4} (7 - 2\Delta)_3 \right] \frac{m_2}{m_1} r_1{}^4 - \right.$$
$$\left. \frac{L_2 u_2}{3 - u_2} \left[\Delta_3 + \frac{\tau_2}{4} (7 - 2\Delta)_3 \right] \frac{m_1}{m_2} r_2{}^4 \right\}, \tag{40}$$

and s is the observed reflexion coefficient as defined by equation (37). The term s_e (vanishing when $u = 0$) is of the order of r^4, $i.\,e.$, very small compared with s_r which is of the order of r^2. Usually we do not commit an appreciable error in assuming that $\Delta_3 = 1$, and that u and τ have values appropriate for the total radiation. Hence equation (40) reduces to

$$s_e = \frac{45}{32} \left\{ L_1 \frac{m_2}{m_1} r_1{}^4 - L_2 \frac{m_1}{m_2} r_2{}^4 \right\}, \tag{40.1}$$

which, as well as equation (39), involves only quantities that can be determined either from spectroscopic observations or from analysis of the light curves within eclipses—except the temperature ratio occurring in equation (39) in the fourth power. It is immediately obvious that the observed reflexion coefficients provide us with a possibility of determining directly the ratios of the temperatures of the two components.

This value has so far been obtainable only indirectly. The usual procedure has been to adopt a definite temperature for the brighter component according to its spectrum (if known), and to derive the

temperature of the secondary from the ratio of the surface brightnesses, with the aid of Planck's formula. This presumes knowledge of the effective wave-length and temperature scale. The more direct way of determining the temperature ratios from equation (39) is probably to be preferred; but if the black-body law holds good and the adopted temperature scale is sound, the ratios arrived at in either way should agree.

The following two tables illustrate the point. TABLE 3 contains the summary of the observational data. Its successive columns indicate:

TABLE 3

(1) Star	(2) Sp$_1$	(3) u_1	(4) L_1	(5) r_1	(6) r_2	(7) Authority
SZ Cam	Bo	0.10	0.72	0.36	0.26	Wesselink, Diss. Leiden. 1938
V Pup	B1	0.06	0.59	0.35	0.34	v. Gent, B.A.N. 317. 1939
VV Ori	B2	0.14	0.90	0.38	0.18	Hertzsprung, Potsd. Publ. 25: 5. 1919.
TT Aur	B3	0.11	0.68	0.35	0.31	Joy-Sitterly, Ap. J. 73: 77. 1931.
u Her	B3	0.12	0.67	0.29	0.30	Dufay, Lyon. Bull. 12: 87. 1930.
u Her	B3	0.19	0.71	0.29	0.30	Baker, Lick Bull. 12: 130. 1926.
μ^1 Sco	B3	0.19	0.57	0.33	0.36	Rudnick-Elvey, Ap. J. 87: 553. 1938.
U Oph	B5	0.16	0.54	0.24	0.24	Shapley, Princ. Contr. 3. 1915.
TV Cas	B9	0.33	0.85	0.28	0.28	McDiarmid, Princ. Contr. 7. 1924.
U Cep	Ao	0.44	0.84	0.20	0.32	Dugan, Princ. Contr. 5. 1920.
RV Oph	Ao	0.44	0.83	0.13	0.20	Dugan, Princ. Contr. 4. 1916.
W UMi	Ao	0.44	0.90	0.37	0.32	Dugan, Princ. Contr. 10. 1930.
RZ Cas	A2	0.54	0.90	0.25	0.29	Dugan, Princ. Contr. 4. 1916.
SZ Her	A3	0.57	0.80	0.33	0.34	Dugan, Princ. Contr. 6. 1924.
X Tri	A3	0.57	0.87	0.29	0.34	Dugan, Princ. Contr. 8. 1928.
Z Dra	A5	0.61	0.91	0.24	0.26	Dugan, Princ. Contr. 2. 1912.†
R CMa	Fo	0.67	0.93	0.25	0.24	Dugan, Princ. Contr. 6. 1924.
RT Per	F2	0.67	0.87	0.31	0.27	Dugan, Princ. Contr. 1. 1911.†

† As reduced by Shapley, Princ. Contr. 3. 1915.

(1) the star, (2) the spectrum of the primary component, (3) the theoretical degree of limb darkening of the primary, taken as function of its effective temperature and wave length from Pannekoek,[†] (4) the luminosity of the primary (total luminosity of the system taken as unity), (5) and (6) the radii r_1 and r_2 (in case of ellipsoidal stars, the geometric mean of the semi-axes), and (7) the authority. If, as in the

† M. N. 95: 733. 1935.

majority of cases, only "uniform" and "completely darkened" solutions had been derived by the computer, the adopted radii and luminosities were obtained by interpolation between the two limiting sets of elements according to the theoretical degree of darkening given in column (3).

Individual columns of TABLE 4 indicate: (1) the star, (2) the ratio of surface brightnesses of the dark sides, (3) the effective wave length,[†]

<center>TABLE 4</center>

(1) Star	(2) J_1/J_2	(3) λ_{eff}	(4) $\log T_{2e}$	(5) $2s$	(6) $(T_1/T_2)_e$	(7) $(T_1/T_2)_r$	(8) $(T_1/T_2)_e'$
SZ Cam	1.26	4230A	4.34	0.001±0.001(p.e.)	1.16	1.03	1.09
V Pup	1.36	5604	4.29	0.001±0.004	1.18	1.04	1.11
VV Ori	1.86	4330	4.20	0.018±0.003	1.28	1.19	1.27
TT Aur	1.68	5290	4.17	0.049	1.28	1.15	1.29
u Her	2.18	5200	4.13	0.017	1.38	1.19	1.29
u Her	2.48	4500	4.13	0.018±0.001	1.38	1.19	1.31
μ^1 Sco	1.64	4500	4.18	0.035	1.22	1.14	1.22
U Oph	1.15	5290	4.17	0.013±0.005†	1.06	1.10	1.11
RZ Cas	32.1	5290	3.67	0.059±0.008	2.05	2.11	2.55
TV Cas	13.6	5290	3.78	0.074±0.008	1.91	1.76	2.09
U Cep	18.4	5290	3.74	0.043±0.010	1.95	1.75	2.06
RV Oph	13.5	5290	3.76	0.014±0.010	1.85	1.84	1.93
W UMi	9.6	5290	3.79	0.038	1.74	1.44	1.90
SZ Her	5.1	5290	3.80	0.030	1.47	1.32	1.52
X Tri	9.9	5290	3.75	0.018	1.66	1.40	1.67
Z Dra	25.3	5290	3.66	0.040±0.006‡	1.85	1.84	1.93
R CMa	17.0	5290	3.62	0.015	1.71	1.58	1.84
RT Per	5.8	5290	3.68	0.022±0.009‡	1.48	1.38	1.57

† As reduced from Shapley's material by Kopal.
‡ As reduced from Dugan's material by Russell & Shapley, Ap. J. 39: 405. 1914.

(4) the logarithm of the effective temperature of the secondary (that of the primary being assumed according to Kuiper's scale[‡] as computed from J_1/J_2 with the aid of Planck's formula), (5) the observed reflexion constants $2s$ with their probable errors (when known), (6) their ratio T_1/T_2, and (7) the ratio of temperatures as determined from the reflexion effect with the aid of equation (39).

A glance at columns (6) and (7) shows that the values of the latter are in general significantly smaller and nearer to unity. The reason is not difficult to find. The values of column (6) are the ratios of

† Unless specified by the observer, we adopt: λ_{vis} = 5290 Å and λ_{pg} = 4250 Å (Unsöld, "Physik der Sternatmosphären," pp. 52–55).
‡ Ap. J. 88: 429. 1938.

proper temperatures of the two stars, while those of column (7) are the ratios of temperatures on the illuminated hemispheres. The latter are generally somewhat different, on account of the heating effect of the incident radiation discussed above. With the aid of equation (36) the ratios of temperatures on the brighter sides can be reduced to the proper temperature ratios (see column (8)).

A comparison of the final temperature ratios arrived at in two independent ways, namely, from the reflexion effect and from the
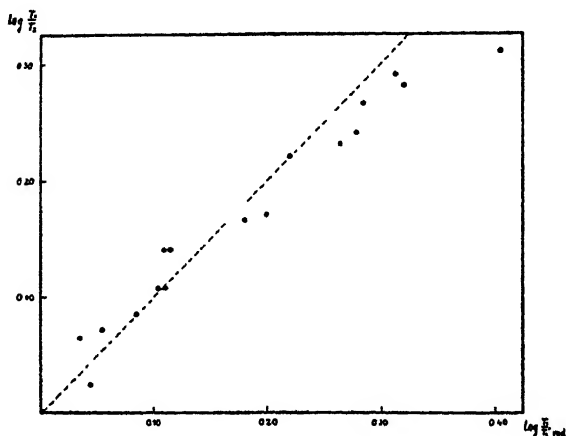

FIGURE 1.

ratios of surface intensities alone, is exhibited diagrammatically in FIGURE 1. As is evident, the agreement is very satisfactory. Exception is noted only in the behavior of some Algol systems with late type secondaries which seem to deviate significantly from the straight line. These deviations are probably real and due to the departures of the components from black-body radiation, the departures affecting both methods differently. The present observational material is admittedly too scarce to permit definite quantitative conclusions. But it is worth mentioning that the deviation of stellar radiation from that of a black body (i. e., the differences between the "effective" and "radiation" temperatures) can be detected from the study of the reflexion effect alone.

Ellipticity Effect

The interpretation of the coefficient of $\cos^2 \psi$ offers relatively more interest, as it depends on the internal structure and the degree of limb darkening in a much more pronounced way than the previous

coefficient. In equation (22) we found that the weighted ellipticity term due to the distorted form of the primary component is $-L_1c_2$, and a similar term arising from the secondary is obtained by interchanging indices. Further, we found that the reflexion effect gives rise also to an appreciable term varying as $\cos^2 \psi$. Combining equations (22) and (30) we can put

$$z = (z_e + z_e') \sin^2 i - z_r \sin i \tag{41}$$

where

$$z_e = \frac{3}{10} \sum_{i=1}^{2} L_i \frac{15 + u_i}{3 - u_i} \left[\Delta_2 + \frac{\tau_i}{4}(5 - \Delta_2) \right] \frac{m_j}{m_i} r_i^3 ; \tag{42}$$

$$z_e' = \frac{15}{8} \sum_{i=1}^{2} L_i \frac{1 - u_i}{3 - u_1} \left[\Delta_4 + \frac{\tau_i}{4}(9 - 3\Delta_4) \right] \frac{m_j}{m_i} r_i^5 ; \tag{43}$$

$$z_r = \sum_{i=1}^{2} L_i \left(\frac{T_j}{T_i} \right)^4 r_i^2 \{ 0\cdot0925 + 0\cdot375 r_j + 0\cdot4698 r_i^2 + \ldots \} \sin i ; \tag{44}$$

and z is the observed ellipticity coefficient from equation (37). The quantities L, m_1/m_2, r and i as well as τ and T_1/T_2 being known, we are in a position, from the observed values of z, to calculate the effects of limb darkening or internal structure. So long as we have observations in only one wave length, we cannot determine these effects for each component separately, and we have to assume that both effects are the same for either component. If this were not so, what we can get is clearly but their mean value, weighted again according to the luminosities of the individual components.

HISTORY OF THE SUBJECT

Of the two quantities u and Δ, the latter is undoubtedly of higher astrophysical interest. The possibility of obtaining in this way some indications concerning the degrees of stellar density condensation was mentioned probably for the first time by Shapley.[*] In his "Study of the Orbits of Eclipsing Binaries" Shapley pointed out, as far back as 1915, that the "uniform values of oblatenesses agree remarkably well with the theoretical value for a homogeneous fluid."[**] In 1930 Russell and Dugan found that the "uniform" ellipticity of the system β Aurigae as determined by Stebbins would correspond to a nearly homogeneous structure.[†] Since this time the question has been the subject of three specific investigations by Walter,[‡] Kopal[§] and

[*] Princ. Contr. 3. 1915.
[**] Princ. Contr. 3: 114. 1915.
[†] M. N. 91: 212. 1931.
[‡] Königsb. Veröff. ff. 2. 1931.
[§] M N. 96: 854. 1936.

Luyten.[†] Apart from some minor differences, all were unanimous in pointing out that stellar density condensations as derived from the ellipticities of close binary systems seem to be very weak. In the investigations by Walter and Kopal the gravitational effect was not considered. Luyten allowed for it in a way which failed to reproduce it correctly.[‡] In the present discussion, formulae which allow strictly for the gravity effect and for intermediate degrees of darkening will be applied for the first time to individual stars.

SCOPE OF THE PRESENT INVESTIGATION

The observational material to which we shall confine our attention consists primarily of a group of B-type eclipsing systems. There are several reasons for this choice. First, close pairs displaying appreciable ellipticities are relatively frequent among B stars, a number of which are also apparently bright enough to be accessible to present-day photoelectric equipment. Secondly, in the B-type systems the components are as a rule very similar in their physical characteristics, which makes the necessary simplifying assumption about the similarity of both components in internal structure acceptable. For the same reason, two W Ursae Majoris-type were also included, in order to extend the investigation to some representative dwarf systems.

As regards the B-type stars, we have in addition a relatively definite idea about their limb darkening. There is good theoretical evidence that the degree of darkening decreases rapidly with increasing temperature, and for the B and O-type stars becomes very small. Quantitative results were recently published by Pannekoek[§] which enable one to find out the degree of darkening corresponding to any effective temperature and wave length. Pannekoek's values may not yet be the last word, but as they are numerically small for the majority of B stars, the errors possibly introduced by their inaccuracy are well within the limits of error of the other observable quantities.

TABLE 5 reviews the material on which the present investigation is based. Successive columns indicate: (1) the star, (2) the source of data, (3) effective wave length, (4) the total number of observations

---

[†] M. N. **98**: 459. 1938.

[‡] Luyten puts (M. N. **98**: 461. 1938), approximately, $L\left(\frac{\pi}{2}, \frac{\pi}{2}\right)/L\left(0, \frac{\pi}{2}\right) = 1 + \frac{3}{2}$ $(1.5 + \alpha)\varepsilon^2$, where ε^2 is the dynamical ellipticity and α is a constant assuming, for Roche's model, the value $\alpha_R = 0.5$, while for Maclaurin's model $\alpha_M = 0.2$. As long as we bear in mind the effects of the total radiation, Luyten's value for α_R is true. His α_M, however, is incorrect, the true value being $\alpha_M = -0.25$.

[§] M. N. **95**: 733. 1935.

N, (5) the number of normal points n, (6) the probable error of a single normal ε_0 (in magnitudes), and (7) the relative weight of the light curve defined as $10^{-5}n/\varepsilon_0^2$.

TABLE 6 contains the elements of our systems. Its individual columns give: (1) the system, (2) the spectrum of the primary, (3) the ratios of surface brightnesses J_1/J_2, (4) the ellipticity constants z with their probable errors, (5) the radii-ratios $k = r_1/r_2$ (the suffix 1 refers to the more massive component in every case), (6) the luminosity of the primary L_1 (total luminosity of the system taken as unity), (7) the semi-major axis of the primary a_1 (expressed in terms of the orbital radius as unity), (8) the cosine of the angle of inclination, and (9) the mass-ratio m_2/m_1.

<div align="center">TABLE 5</div>

(1) Star	(2) Authority	(3) λ_{eff}	(4) N	(5) n	(6) ε_0	(7) $10^{-5}n/\varepsilon_0^2$
					m	
SZ Cam	Wesselink, Diss. Leiden. 1938.	4230A	12479	104	0.005	41.6
V Pup	van Gent, B. A. N. 8: 319. 1939.	5604	1105	86	0.009	10.6
RZ Cen	Shapley, Princ. Contr. 3. 1915.	4250	138	35	0.010	3.5
VV Ori	Hertzsprung, Potsd. Publ. 23: 5. 1919.	4330	1451	66	0.006	18.3
σ Aql	Wylie, Ap. J. 56: 232. 1922.	4500	97	37	0.006	18.3
u Her	Baker, Lick Bull. 12: 130. 1926.	4500	161	57	0.005	22 0
μ^1 Sco	Rudnick-Elvey, Ap. J. 87: 553. 1938.	4500	128	20	0.004	12.5
U Oph	Shapley, Princ. Contr. 3. 1915.	5290	266	53	0.015	3.4
S Ant	Joy, Ap. J. 64: 288. 1926.	5290	†	23	†	†
W UMa	Schilt, B. A. N. 3: 20. 1925.	4250	468	42	†	†
	Walter, Königsb. Beob. 45: 1. 1928.	4250	477	48	0.012	3.4

† Data not available.

<div align="center">TABLE 6</div>

(1) Star	(2) Sp_1	(3) J_1/J_2	(4) $2z$	(5) k	(6) L_1	(7) a_1	(8) cos i	(9) m_2/m_1
SZ Cam	B0	1.26	0.117±0.002 (p.e.)	1.41	0.72	0.39	0.339	0.63
V Pup	B1	1.36	0.277±0.010	1.02	0.59	0.38	0.226	0.77
RZ Cen	B2	0.7	0.188	2.00	0.74	0.49	0.210	0.50
VV Ori	B2	1.86	0.081±0.007	2.22	0.90	0.40	0.000	0.34
σ Aql	B3	1.34	0.075±0.010	1.00	0.57	0.24	0.312	0.83
u Her	B3	2.48	0.152±0.006	0.96	0.71	0.31	0.237	0.38
μ^1 Sco	B3	1.64	0.169	0.90	0.57	0.36	0.470	0.83
U Oph	B5	1.15	0.081±0.023	1.00	0.54	0.25	0.075	0.88
S Ant	A8	1.2	0.340	1.28	0.67	0.50	0.468	0.56
W UMa	F8	1.1	0.415±0.010	1.15	0.59	0.45	0.174	0.72

REMARKS TO TABLES 5 AND 6

SZ CAMELOPARDALIS.—Wesselink's light curve of this variable is probably one of the most precise ever made for an eclipsing variable. Yet despite this, two solutions are derived by Wesselink with $k = 1.4$ and 1.0, which represent the observations equally well; and in the absence of spectroscopic observations of both components the solution might seem indeterminate. But using Plaskett's few radial velocity measurements, Wesselink approximated the mass-function as

$$\frac{m_2^3 \sin^3 i}{(m_1 + m_2)^2} = 0.475 \; \odot.$$

The value of i differs but slightly for both solutions, therefore we can put in either case

$$\frac{m_1}{m_2} = 1.20 \; (m_1 + m_2)^{1/3} - 1,$$

which yields,

$(m_1 + m_2)$	m_2/m_1
5 $\odot$	0.95
10	0.63
15	0.51
20	0.44
25	0.40

The combined mass of a B0 system is likely to be fairly high; consequently the mass-ratio must be around 1.5 or greater—which would be at variance with the first solution ($k = 1.0$), for in this case the components should have nearly the same luminosity ($L_1 = 0.55$), and for $m_1/m_2 \sim 1.5$ the companion would have to deviate largely from the mass-luminosity law. The second solution ($k = 1.4$) gives much better agreement. If we assume that both components fit the empirical mass-luminosity law, their mass-ratio should be approximately $m_1/m_2 = 1.6$, and the combined mass about 10 $\odot$.

V PUPPIS.—The photometric solution is taken from van Gent. The ratio L_1/L_2 is spectroscopic (Maury, H. A. 84: 172). The spectrograms show lines of both components, but the measurements are relative and so only the total mass of this system is known ($m_1 + m_2 = 34.8 \; \odot$). The luminosity criterion yields $m_1/m_2 = 1.3$.

RZ CENTAURI.—Orbital elements are taken from Shapley's Catalogue (Princ. Contr. 3. 1915, "first grade orbit"). Observations are those by Leavitt (H. A. 60: 5). In considering their accuracy Shapley wrote: "The measurements of photographic plates, though few in number, are of relatively high accuracy and the light curve is of unusual value" (op. cit. p. 22).

No radial velocity observations. The luminosity criterion yields $m_1/m_2 = 2$.

VV ORIONIS.—There exist two good light curves of this variable, one photographic by Hertzsprung (op. cit.) and the other photoelectric (rubidium cell) by Schneller (Klein. Veröff. Berlin Babelsberg 17: 1936). Both Hertzsprung and Schneller carried out solutions of their light curves; in Hertzsprung's case an inclination of 90° was assumed. Unfortunately, contrary to Hertzsprung, Schneller did not give the details of his solution, nor the probable errors of his results. Consequently the writer adopted the geometrical elements as derived by Hertzsprung. The eclipses are annular and total in turn, and the fair agreement between the photographic and photoelectric orbits suggests that our present picture of this system is fairly reliable.

Spectroscopic observations of both components owing to the faintness of the secondary are lacking. The mass-function of VV Orionis according to Daniel (Allegh. Publ. 3: 179) is

$$\frac{m_2^3 \sin^3 i}{(m_1 + m_2)^2} = 0^{.}358 \odot \; ,$$

which for $i = 90°$ yields

$$\frac{m_1}{m_2} = 1.41 \; (m_1 + m_2)^{1/3} - 1.$$

We find that

$(m_1 + m_2)$	m_2/m_1
10 $\odot$	0.49
15	0.40
20	0.35
25	0.32
30	0.29

The luminosity criterion yields $m_2/m_1 = 0.29$. Plummer (M. N. **98**: 137. 1938) from statistical considerations finds $m_2/m_1 = 0.39$. We adopt the mean $m_2/m_1 = 0.34$, leading to a rather high combined mass of 23 O.

VV Orionis is triple, the eclipsing system being attended by a third body revolving in a period of 120 days (Daniel, *op. cit.*). Its mass, however, can be but very small and its luminosity will scarcely amount to more than one per cent of the total luminosity of the system.

σ AQUILAE.—The ratio L_1/L_2 is spectroscopic, as well as the mass-ratio $m_2/m_1 = 0.83$ (Wylie, *op. cit.*).

u HERCULIS.—This is undoubtedly the best known system included in our discussion. In addition to a photometric light curve by Dufay (Lyon Bull. **12**: 87. 1930), there are three photoelectric light curves observed with a potassium cell by Pavel (A. N. **217**: 177. 1922), Baker (Lick Bull. **12**: 130. 1926), and Shapley-Calder (Harv. Circ. 425. 1937), in addition to one photographic light curve by Baker (Laws Bull. **12**: 209. 1921)—all being in mutually excellent agreement. The solutions by Dufay ($\lambda_{eff} = 5200$ Å) and Baker (ensemble of photoelectric observations by Pavel and Baker) agree so closely that there is but little room for doubt concerning the presently adopted geometrical elements.

The mass-ratio $m_2/m_1 = 0.38$ is spectroscopic (Baker, Allegh. Publ. **1**: 77).

μ^1 SCORPII.—The observations were made with the aid of a potassium cell. The ratio L_1/L_2 is spectroscopic (Maury, H. A. **84**: 172). Only relative radial velocity orbit known, leading to $m_1 + m_2 = 24$ O. The luminosity criterion yields $m_1/m_2 = 1.2$.

U OPHIUCHI.—Observations were made with the polarizing photometer by Pickering and Wendell (H. A. **46**: 212. 1882; also Wendell, H. A. **69**: 1. 1902), each observation being the mean of twelve sets. The orbital elements are taken from Shapley (Princ. Contr. 3, "first grade orbit"). The ellipticity and reflexion constants have been re-determined for the purposes of the present investigation by least squares, with the use of Shapley's normals; as a result, z comes out slightly smaller than that given by Shapley; but no change in geometrical elements appears necessary.

The mass-ratio $m_2/m_1 = 0.88$ is spectroscopic (Plaskett, D. A. O. **1**: 140. 1919).

S ANTLIAE.—Photometric observations are those by Wendell. The mass-ratio $m_2/m_1 = 0.56$ is spectroscopic (Joy, *op. cit.*).

W URSAE MAJORIS.—There exist two photographic orbits of nearly equal weight by Schilt and Walter, the mean of which is adopted. The mass-ratio $m_2/m_1 = 0.72$ is spectroscopic (Adams & Joy, Ap. J. **49**: 189. 1919).

GENERAL CHARACTERISTICS

All the systems listed in the above tables have several characteristics in common. A glance at the column containing the ellipticity constants z shows that there is no inherent difficulty in their determination—it is merely a matter of the accuracy of the observations. For the best of our light curves, that of SZ Camelopardalis, z has been determined by Wesselink with a probable error of about two per cent, and for an average star of TABLE 6 it amounts to less than 10 per cent. The same is true for the ratios of surface brightnesses J_1/J_2. Unfortunately, the relative dimensions of the components can be derived with a much lower degree of accuracy. Apart from VV Orionis where the primary minima happen to be annular, and RZ Centauri with grazing eclipses, all our stars are partially eclipsing systems, and all investigators are unanimous in pointing out that the determinateness of solutions rapidly diminishes with decreasing depth of minima. For this reason our knowledge of a system such as, say, u Herculis can be regarded as fairly reliable; but for the rest of our stars the

solutions are intrinsically indeterminate over a certain range of k—no matter how accurate the light curve. Fortunately, spectroscopic observations may prove to be of some assistance. The radial velocity amplitudes of the B systems are usually so large that the lines of both components are easily separated. The intensities of the lines of each component can give us a good idea about their relative luminosities. Their radii-ratios can then be determined from the well-known relation $L_1/L_2 = k^2 J_1/J_2$. The radii-ratios of V Puppis, σ Aquilae, and μ^1 Scorpii have been determined in this way. Such determinations, although probably better than those based on the form of the light curve alone, cannot claim any particular accuracy;[†] neither can the values of a and i determined subsequently from the light curve with the use of fixed k.

This is a troublesome fact, because the radii-ratios of the components occur in equation (42) in the third power. We have consequently to expect that the principal source of uncertainty in deriving the density condensations will not be due, say, to the inaccuracy of the observed ellipticity constants, but to the inadequate precision with which the radii of both components can be determined. Indeed, in analyzing equation (42) we find that (apart from the effects of darkening)

$$\frac{\partial f}{\partial L_i} = \frac{z_e}{3} \left\{ \frac{m_1}{m_2} r_2{}^3 - \frac{m_2}{m_1} r_1{}^3 \right\} Q^2, \tag{45.1}$$

$$\frac{\partial f}{\partial \left(\frac{m_i}{m_j}\right)} = \frac{z_e}{3} \frac{m_j}{m_i} \left\{ L_1 \frac{m_2}{m_1} r_1{}^3 - L_2 \frac{m_1}{m_2} r_2{}^3 \right\} Q^2, \tag{45.2}$$

$$\frac{\partial f}{\partial r_i} = - z_e L_i \frac{m_j}{m_i} r_i{}^2 Q^2, \tag{45.3}$$

where

$$f = \left[\Delta_2 + \frac{\tau_1 + \tau_2}{8} (5 - \Delta_2) \right]$$

and

$$\frac{1}{Q} = \sum_{i=1}^{2} L_i \frac{m_j}{m_i} r_i{}^3.$$

We see that errors arising from the uncertainty of the observational data are minimal, if k, J_1/J_2, and m_1/m_2 are equal to unity. Equations (45.1, 2) then reduce to

---

[†] Chiefly because the line intensities are expected to display periodic variations due to the reflexion effect—a fact observed by Miss Maury (H. A. 86: 157) and recognized as such by Struve (Ap. J. 86: 198).

$$\frac{\partial f}{\partial L_i} = \frac{\partial f}{\partial \left(\frac{m_i}{m_j}\right)} = 0, \qquad \text{but} \qquad \frac{\partial f}{\partial r_1} = -\frac{1}{2}\frac{z_e}{r_i^4}; \qquad (45.4)$$

i. e., errors in r will sensibly affect the accuracy of Δ_2 calculated from equation (42) under any circumstances.

The last observational quantity entering (42) is the mass-ratio of the two components. In the two-spectra systems this can be derived from the observed amplitudes of the radial velocity changes.[†] But if the orbit of only one component is known, or if radial velocity measurements are completely lacking, recourse must be made to indirect methods depending on the characteristics of each individual system. The mass-ratios determined indirectly by the use of statistical methods are naturally more liable to be in error. However, we have just seen that as long as the masses of both components are not too unequal, the errors in the mass-ratios do not affect the Δ's too sensibly.

RESULTS

The final step of our discussion is the calculation of the constants Δ_2 specifying the degrees of density condensation. This proceeds by use of equations (42–44). The radii r of spheres having the same volume as actually distorted ellipsoids are approximately

$$r_i = a_i\left(1 - \frac{1}{6}\left(1 + \frac{7m_j}{m_i}\right)\Delta_2 r_i^3 + \dots\right) \qquad (46)$$

We note that in working with the third and higher powers of r we may, within the scheme of our accuracy, put $r_i = a_i$.

The determination of the Δ_2's is then straightforward. The "effective" coefficients of limb darkening occurring in equation (22) would be identical with Pannekoek's theoretical values,[‡] were it not for the blanketing effect of the reversing layer. If a star exhibits an absorption spectrum, a part of the photospheric radiation dammed back in the Fraunhofer lines will be re-distributed in direction, and the rate of center to limb variation of intensity will be reduced. The "effec-

[†] To be accurate, the ordinary spectroscopic mass-ratios should still have added to them a correction for the reflexion effect. Owing to it, the centers of light on the two discs do not project themselves precisely on the centers of mass, but are displaced toward the center of gravity. Consequently, the measured radial velocity amplitudes are systematically underestimated. The amount of appropriate correction (Pike, Ap. J. **73**: 205. 1931; or Kuiper, Ap. J. **88**: 498. 1938) is, however, very small in most practical cases, and its application in the present paper might give an improper idea as to our degree of accuracy. Therefore it was not considered.

[‡] M. N. **95**: 733. 1935.

tive" coefficient of limb darkening is then related to the true coefficient u_0 by[†]

$$\frac{u}{u_0} = \frac{5 - 5\eta}{5 - \eta},$$

where eta is the fraction of light returned by the reversing layer to the photosphere. Its value depends chiefly on the spectral type. For our B-systems, η is likely to be too small to be significant; for S Antliae and W Ursae Majoris, however, we shall adopt for it the value empirically determined for the Sun ($\eta_\odot$ = 0.08, according to Mulders, Diss. Utrecht 1934).

The values of τ can be taken from TABLE 2. Having corrected the observed ellipticity coefficients for the reflexion effect and orbital inclination, we can determine the Δ's. In doing so we note that (TABLE 1) Δ_4 is in every case very near unity. Thus, as a first approximation, we may put $\Delta_4 = 1$ and evaluate Δ_2; then we take from TABLE 1 the corresponding Δ_4 and repeat the calculation of Δ_2. Usually the second approximation is amply sufficient. The Δ_2's arrived at in this way are the mean values of the true Δ_2's corresponding to the individual components; but our systems were selected so as to make the discrepancies between the individual and mean Δ_2's as unlikely as possible. The resulting Δ_2's for our ten systems and their estimated probable errors are given in TABLE 7.

TABLE 7

Star	Δ_2	wt
SZ Cam	1 03±0 06 (p.e.)	3
V Pup	1 4 ±0 2	1
RZ Cen	0 9 ±0 2	1
VV Ori	0 9 ±0 2	1
σ Aql	1 2 ±0 3	1
u Her	1 03±0 04	5
μ^1 Sco	1 08±0 08	2
U Oph	1 1 ±0 2	1
S Ant	0 8 ±0 3	1
W UMa	0 9 ±0 2	1

DISCUSSION OF RESULTS

In comparing the above Δ_2's with the theoretical values corresponding to the various degrees of density condensation given in TABLE 1, we see at once that the empirical Δ_2's are in accordance with the view

[†] Milne, The Observatory **51**: 88. 1928.

recently advanced, also from the study of apsidal motion,[†] namely that the density condensations of real stars are high. The weighted mean of the above Δ_2's is

$$\overline{\Delta}_2 = 1.03 \pm 0.03 \text{ (p. e.)},$$

which shows that if stars are polytropes, their structure seems to be specified very nearly by the index $n = 3$. The general feature of equation (6) shows that with increasing central condensation Δ_2 rapidly approaches unity. The present result is therefore inadequate as an upper limit to the density condensations of our systems, which are well distributed along most of the main sequence. However, it definitely imposes a lower limit to possible condensations of genuine main-sequence stars; on the basis of our results it appears improbable that these condensations are lower than, say, that of a polytrope with $n = 2.5$.

A further result of the present discussion is the indirect check on the reality of the gravitational effect. This has not been so obvious as it may seem. The theoretical proof given by von Zeipel and Chandrasekhar depends on the assumption that there is no internal circulation of matter. A system of rotating currents, for example, would tend to equalize the temperature gradient beneath the photosphere, and could eventually be effective enough to make the temperature gradient independent of direction. The gravity effect would then disappear, and (apart from the effect of limb darkening) the optical ellipticity would be identical with the dynamical one. We have seen before that for centrally condensed configurations the theoretical gravity effect makes the optical ellipticity twice the dynamical. From our observational material we were able to derive the mean Δ_2 with an uncertainty of about 3 per cent—and its agreement with general theoretical expectations, as well as with results based on the study of observed apsidal motions, strongly suggests that it is very near to reality. We can conclude, therefore, that the observed gravity effect amounts to at least 95 per cent of its theoretical value; *i. e.*, the stars cannot possess any system of rotating currents below the photosphere capable of influencing the direction-distribution of the temperature gradient to any appreciable extent.

MASSIVE STARS

The statement about the compatibility of actual stellar structures with the centrally condensed model is, perhaps, subject to one exemp-

[†] Sterne, M. N. **99**: 662. 1939.

tion: that relating to massive stars. Among our present material we find two fairly massive systems, μ^1 Scorpii and V Puppis, and especially for the last one Δ_2 seems to deviate significantly from the mean. In addition, another very massive eclipsing system, UW Canis Majoris, exhibits marked ellipticity and lends itself to an extension of the proof. Rudnick and Elvey obtained recently a photoelectric light curve of this system, and derived the orbital elements. Their values as published so far[t] are only provisional and therefore this system was not included in TABLES 5, 6, and 7. It is, however, tempting to adopt their provisional elements for the purpose of deriving Δ_2. We obtain $\Delta_2 = 2.3$. It is difficult to estimate its probable error, as the original light curve as well as the details of the orbital solution are not yet accessible; but it is unlikely that the uncertainty exceeds, say, ± 0.5. This would make the structure of UW Canis Majoris definitely incompatible with that of moderately massive systems and its central condensation would only be weak.

The following tabulation will illustrate the point. The second column gives the total mass of the system; column (3) the Δ_2's; column (4) the limits in polytropic class derived from the probable errors of Δ_2; column (5) the most probable value of the central condensation ρ_c/ρ_m, provided the configurations are polytropes.

TABLE 8

(1) Star	(2) Mass	(3) Δ_2	(4) Polytropic index	(5) ρ_c/ρ_m
μ^1 Sco	24.0 ⊙	1.08±0.08	5>n>2	33
V Pup	34.8	1.4 ±0.2	2>n>1	5.3
UW CMa	77.7	2.3(±0.5)	0.7>n>0	2.0

The mass-ratios in all the three systems are near unity; consequently the masses of the individual components will be of the order of one half of the combined mass. This shows, therefore, that starting with masses of about 10 ⊙ actual stellar structures begin to deviate from the centrally condensed model. The density condensations, which for stars with $m < 10$ ⊙ must be high, become gradually less pronounced, until, if extrapolation can be any guide, very massive stars (of the order of 100 ⊙) may approach the limits of homogeneity.

This remarkable feature of massive eclipsing systems, as inferred

[t] Quoted by Kuiper, Ap. J. 88: 503–504. 1938.

from the study of their observable properties, is in complete agreement with recent theoretical considerations by Chandrasekhar[†] based mainly on the theory of the hydrogen content of the stars. The importance of further study of massive eclipsing systems as testing stones for modern astrophysical theories is obvious. From the practical point of view, we may note that considerably more accurate results than the present ones can be expected when radiometric light curves become available. Observations in the total light will provide us with ellipticity coefficients independent of the degree of limb darkening as well as of the effective temperature and wave length, and the relation between the geometrical elements and the internal structure will be much simplified. All of the three, massive systems discussed in this paragraph are easily accessible to present-day radiometric equipment. It is to be hoped that the present results will stimulate interest in their radiometric study—all the more so because, for low density condensations, the ellipticity becomes a very sensitive indicator of the central condensation; and it is the only possible "dynamical" indicator, as, owing to the vanishing eccentricity of their orbits, there is no hope of obtaining alternative indications from the study of the apsidal motion.

Deviations from Equilibrium

The discussion carried out so far is relevant only to systems obeying the equilibrium conditions stated in the introduction to the present paper. The discussion, however, would not be complete without mentioning that there are certain observable phenomena which definitely do not fit within this scheme. These, in brief, are variability of orbital periods, and asymmetry of light curves for circular orbits.

VARIABILITY OF PERIODS

Observations extending over a sufficiently long interval show frequently that the orbital periods of many eclipsing systems are subject to variations which may or may not seem to be of a periodic nature. One way of accounting for the periodic terms is to assume that the eclipsing system is attended by a third body. Its presence would give rise to both real and apparent period changes due to perturbations of the eclipsing orbit, light equation, and rotation of the apsidal line. A test for the existence of a third body can be provided by radial velocity observations. If the velocity of the center of mass does not

† *E.g.* Proc. Amer. Phil. Soc. **81**: 181. 1939.

exhibit variations synchronous with those of the orbital period, the hypothesis of the existence of a third body of appreciable mass must be discarded. The quantities specifying period variations (if periodic) permit a conclusion about the mass of the hypothetical third body even in the absence of spectroscopic observations. The resulting masses sometimes prove to be inconsistent with other physical evidence;[†] the auxiliary hypothesis of a third body has then to be dropped definitely, and the observed period changes must be explained on the basis of a two-body system alone.

Here again we face an alternative: the period changes may either be apparent and due to the rotation of the apsidal line in an eccentric orbit; or they may be real and due to perturbations arising from the distorted form of the two components. The decisive feature is whether or not the primary and secondary minima deviate from ephemeris in the same sense. If they do, this is contrary to what could be accounted for on the basis of the rotation of the apsidal line, and we are led to consider the general perturbations. But here we meet an essential difficulty. In binary systems obeying our equilibrium conditions the orbital periods should be constant. The perturbing function due to enforced oscillations of the fluid components is always radial and cannot account for any perturbations with periods uncommensurable with that of the orbit.[‡]

The following four Algol-type systems can be taken as examples. The sources given below refer to investigations of the period changes.

RS CVn	Sitterly, Princ. Contr. 11, 1930
U Cep	Carrasco, Madrid Bol. 16, 1933
RV Oph	Dugan, Princ. Contr. 4, 1916
RT Per	Dugan, Princ. Contr. 17, 1938

U Cephei has been under nearly constant observation for seventy years, and is one of the best investigated stars in this respect. The presence of a third body seems to be ruled out by the constancy of the radial velocity of the center of mass; in addition, the period changes are of such a complex character that any attempt to explain them on the basis of a third-body hypothesis seems to be *a priori* destined to fail. For RT Persei spectroscopic observations are lacking at the present time, but Dugan's attempt to explain the observed period changes by the presence of a third body has led to unsatisfactory results (*op. cit.*), and the parallel displacement of both minima dis-

† Kopal, A. N **258**: 393. 1936.
‡ The writer owes this remark to Dr. Sterne

poses of the possibility that revolution of the apsidal line can explain any appreciable part of the observed period changes. Similar conclusions apply to RS Canum Venaticorum and RV Ophiuchi.

The four stars mentioned were selected merely to illustrate the point and the list is by no means regarded as complete. On the contrary, in a recent study of the periods of eclipsing binaries Dugan[†] came to the conclusion that "variations of periods have been confirmed in eighteen out of twenty six cases studied, and the published sine and parabolic terms are found to be of only temporary applicability." This suggests that, on the other hand, changing periods may be a general feature of eclipsing binaries, and systems with periods constant over long intervals of time may be comparatively rare.

If we re-consider the equilibrium conditions imposed in the introduction we find that one of them is plainly hypothetical: that about the perpendicularity of the rotational axes of both components to the orbital plane. It is possible that the period changes observed in a number of systems are perturbations arising from the break-down of this last condition. The deviations from perpendicularity may be only small and variable in time, or may apply to only one component; but the orbital periods of eclipsing variables can be determined with such a degree of precision that even very small perturbations can be detected in this way.[‡]

ASYMMETRY OF LIGHT CURVES

It is very significant in this connection that systems exhibiting period changes frequently display also asymmetric light curves. The asymmetry may, as is well-known, be produced also by the orbital eccentricity. But the eccentricity produces generally three observable phenomena, namely, asymmetry of light curve, displacement of secondary minimum, and unequal duration of both minima. These three phenomena do not need to occur simultaneously; yet it is essential that the asymmetry can never occur without the secondary minimum being displaced and, generally, both minima being of unequal duration. Consequently, if the two latter effects of orbital eccentricity (which can be observed much easier than the asymmetry) are

[†] A. J. **46**: 148. 1937.

[‡] In the case of precessional motion of the rotation axes the change in inclination changes also the distribution of brightness over the visible hemisphere (the gravity effect being symmetrical with respect to the axis of rotation) and hence should affect also the optical ellipticity. But the possibility of detecting changes in inclination in this way seems to be only of theoretical interest; in practice, the effect will be too small to be observable.

(a) by the number of atoms per gram: $1/m_H A$; A is the atomic weight, m_H the mass of hydrogen atom;

(b) by the probability of finding a given state of energy E occupied, *i. e.*,

$$f(E) = \frac{1}{e^{(E+\Psi)/kT}+1}$$

where $(\Psi/kT)^*$ characterizes the "degree" of degeneracy of a Fermi-Dirac gas ($\Psi/kT > 0$ implies non-degeneracy, $\Psi/kT < 0$ implies degeneracy);

(c) by the probability of finding a state of energy $E + h\nu$ empty, *i. e.*,

$$1 - f(E + h\nu) = \frac{1}{1 + e^{-(E+h\nu+\Psi)/kT}}$$

and (d) by integrating the product over E.

If this process is carried out and the integration is performed between the limits zero and infinity, we get for the absorption coefficient due to free-free transitions[†] the following:

$$(3) \quad a_\nu = \left(\frac{32\pi^2 Z^2 e^6 kT}{3\sqrt{3}ch^4 A\, m_H}\right) \cdot \frac{1}{\nu^3} \cdot \frac{e^{h\nu/kT}}{e^{h\nu/kT} - 1} \cdot \log_e\left[\frac{e^{h\nu/kT}(1 + e^{\Psi/kT})}{(1 + e^{(h\nu+\Psi)/kT})}\right]$$

Under certain conditions, especially in the case of high degeneracy, these "free-free" transitions give the largest contribution to the absorption coefficient. In this case, the opacity can be calculated simply by inserting (3) in the Rosseland mean (1). The log term prevents an exact integration of (1) and some approximation must be made, depending on the value of Ψ/kT. On the non-degenerate side ($\Psi/kT > 0$) we can get by employing a suitably chosen series expansion for the log term, the following rapidly convergent series for K_R:

$$(4) \quad \frac{1}{K_R^{N.D.}} = \frac{15 \cdot 7! \xi_7}{4\pi^4 \Theta}\left(\frac{kT}{h}\right)^3 \cdot \frac{1}{\log_e(1+e^{-\Psi/kT})} \cdot \left\{1 + \right.$$

$$\left. \frac{[2\log_e(1+e^{-\Psi/kT})+e^{-\Psi/kT}]}{256\log_e(1+e^{-\Psi/kT})} + \cdots\right\}$$

where we have taken $\Theta = \dfrac{32\pi^2 Z^2 e^6 kT}{3\sqrt{3}ch^4 A\, m_H}$ and $\xi_7 = \displaystyle\sum_{q=1}^{\infty}\frac{1}{q^7} = 1.0083$.

On the degenerate side ($\Psi/kT < 0$) it is difficult to find a suitable approximation by means of series expansions. However, the integral

---

[*] $\Psi = -J$ where J is the usual Fermi energy.

[†] This limitation to free-free transitions will be removed below and the expression generalized to include bound-free transitions.

(5)
$$\int_0^\infty \frac{\nu^7 e^{h\nu/kT}}{(e^{h\nu/kT}-1)^2 \log_e \left[\dfrac{e^{h\nu/kT}(1+e^{+\Psi/kT})}{1+e^{(h\nu+\Psi)/kT}} \right]} d\nu$$

has a fairly sharp maximum at $h\nu/kT = 7$ so that the slowly varying log term may be replaced by its value for $h\nu/kT = 7$ and taken outside the integral.* We thus find for the degenerate radiative opacity:

(6)
$$\frac{1}{K_R\text{D.}} = \frac{15\cdot 7!\xi_7}{4\pi^4\Theta} \left(\frac{kT}{h} \right)^3 \cdot \frac{1}{\log_e \left[\dfrac{(1+e^{-\Psi/kT})}{(1+e^{(-\Psi/kT-7)})} \right]}$$

This expression may, to the same degree of approximation, be used on the non-degenerate side. In the limit $\Psi/kT \to \infty$, equation (6) goes into the ordinary Kramers opacity law, and in the limit $\Psi/kT \to -\infty$ it becomes the usual expression given for the completely degenerate case.[†] That this expression is valid quantum mechanically follows from the correspondence principle.[‡]

To include the bound-free transitions in addition to the free-free transitions, one writes for the radiative opacity:

(7)
$$K_R = \frac{4\pi^4\Theta}{15\cdot 7!\xi_7} \left(\frac{h}{kT} \right)^3 \log_e \left(\frac{1+e^{-\Psi/kT}}{1+e^{-(\Psi/kT+7)}} \right) \cdot \frac{g\cdot 196.5}{\tau}$$

This formula differs from (6) only by the factor $\dfrac{g\cdot 196.5}{\tau}$. Here g is the well-known Gaunt-factor[§] which gives the ratio by which the classical expression for the bound-free transition probability must be multiplied in order to give the correct quantum mechanical one. For the various absorption edges it ranges from .8 to 1. It is found that a good over-all approximation for g is .875. The factor τ is the so-called guillotine factor; it is defined by equation (7). This definition gives the ordinary corrected Kramers formula in the limit of non-degeneracy. τ increases in magnitude as the occupation of bound states increases and offsets the increase of opacity with ρ as given by its appearance in Ψ/kT. That is, the increase in K_R is at first propor-

---

* In the region where bound-free transitions dominate, the maximum lies closer to $h\nu/kT = 6$. Cf. the paper by P. M. Morse, Ap. J. 92: 27. 1940. We are indebted to Prof. Morse for permitting us to see his manuscript prior to publication. In the region our formulae apply the results are insensitive to the choice of the maximum.

† Majumdar, R. C. Astron. Nach. 244: 65⁻ 1931.

‡ Cf., also, Menzel, D. H. & Pekeris, C. L. M. N. 96: 77. 1936, where this is shown in detail.

§ Gaunt, J. A. Proc. Roy. Soc. A 126: 654. 1930.

tional to ρ but then lags behind as the shells are more and more filled up.

The guillotine factor τ is computed according to the procedure given by Stromgren.* When the density is small and the temperature high, τ is close to one. In this case the ratio of the total photo-electric absorption to the free-free part is $.875 \times 196.5 \approx 172$. On the other hand when $\tau = 196.5$ the bound-free contribution to K_R has disappeared and only the free-free contribution remains. This occurs at about $\Psi/kT = -5$.

We have calculated the guillotine factor in the transition region $(2 > \Psi/kT > -3)$ for a variety of temperatures and a "Russell" mixture. This mixture consists of:

0 (8/16)	Na, Mg (4/16)	Si (1/16)
K, Ca (1/16)	Fe (2/16)	

The calculation has been performed for the unscreened Coulomb field; but an estimate which has been made of the effect of screening shows that the value of τ is not changed appreciably (about 20% in the region where the radiative contribution to the opacity is important). TABLE 1 gives the results.

TABLE 1

The Guillotine Factor τ						
$\Psi/kT \longrightarrow$ 2 $T \downarrow$	1	0	−1	−2	−3
$2 \cdot 10^6$	3.8					
$4 \cdot 10^6$	6.0	11	27	45		
$6 \cdot 10^6$	6 5	14	29	59	100	
$8 \cdot 10^6$	8.6	15	35	68	108	
$10 \cdot 10^6$	10	17	37	71	113	145
$12 \cdot 10^6$	11	18	38	72	116	150
$14 \cdot 10^6$	10	18	39	74	118	153
$16 \cdot 10^6$	11	18	40	75	119	155
$18 \cdot 10^6$	11	18	40	75	119	156

When the region of extreme degeneracy is reached, that is for $(-\Psi/kT) > 40$, one must multiply the opacity by the factor

$$(7a) \qquad \frac{\sqrt{3}}{\pi}\left[\log_e(-\Psi/kT) - 1.85\right]$$

* Stromgren, B.   Z. f. Ap. 4: 118.   1932.

Expression (7a) arises when account is taken of the fact that the mass absorption coefficient contains a factor $\log_e (- \Psi^o/h\nu)^*$ for $E > h\nu$ (*i. e.* when the maximum of the Planck distribution falls within the Fermi threshold) and the new value of $a\nu$ inserted into equation (1). In the white dwarf stars this modified opacity formula would certainly have to be used were it not for the fact that the conductive opacity dominates the radiative opacity when conditions of strong degeneracy are reached. Hence, for a "Russell" mixture of concentration X_R, the final expression which is convenient to apply is:

$$(7b) \qquad K_R = \frac{6.0 \cdot 10^{17} X_R}{\tau T^2} \log_e \left[\frac{1 + e^{-\Psi/kT}}{1 + e^{-(\Psi/kT + \eta)}} \right]$$

THE CALCULATION OF THE CONDUCTIVE OPACITY

In the interior of stars, particularly dense stars, energy is not transported solely by radiative processes but also by electronic conduction. If the energy current is Q, the thermal conductivity λ of the electrons is defined as: $\lambda = - Q \Big/ \dfrac{dT}{dx}$ with $\dfrac{dT}{dx}$ the gradient temperature T, along the direction x. If we compare this formula with the radiative formula:

$$(8) \qquad \frac{dT}{dr} = - \frac{3}{4aT^3} \cdot \frac{K_R \rho}{C} (L_r/4\pi r^2)$$

where K_R is the radiative opacity and $(L_r/4\pi r^2)$ really the energy current (for a specification of the meaning of the other symbols *cf.* Eddington),[†] we see that the conductive opacity K_C is given by:

$$(9) \qquad K_C = \frac{4acT^3}{3\rho\lambda}$$

It is important to make a fairly accurate calculation of the conductive opacity since it plays an important role in the white dwarf stars. In the transition region between non-degeneracy and degeneracy it contributes as much to the energy transport as does radiation; while in the degenerate core it completely dominates the energy transport so that only the conductive part of the opacity is significant.

We calculate the thermal conductivity λ to a "second" approximation on both the degenerate and non-degenerate sides—in accordance

---

* Chandrasekhar, S. Proc. Roy. Soc. **133**: 241. 1931.
† "Internal Constitution of the Stars." 1930.

with the usual procedure adopted in treating the problem in metals. If $f(\mathbf{v})$ is the electron distribution function in the presence of a temperature gradient $\dfrac{dT}{dx}$ then the particle current j and the energy flow Q are given by:

$$(10) \qquad j = \frac{2m^3}{h^3} \int v_z f(\mathbf{v}) d\mathbf{v}$$

$$(11) \qquad Q = \frac{2m^3}{h^3} \int E v_z f(\mathbf{v}) d\mathbf{v}$$

In these equations E is the energy of the electron, $\mathbf{v}$ its velocity. Since there is no net convection of electrons, we have the condition $j = 0$. This can only be the case when there is an electric field F in addition to the temperature gradient.[*]

The distribution function $f(\mathbf{v})$ is calculated from the Boltzmann stationary condition:

$$(12) \qquad \left(\frac{\partial f}{\partial t}\right)_{\leftarrow \text{temp.-gradient}} + \left(\frac{\partial f}{\partial t}\right)_{\text{electr. field}} + \left(\frac{\partial f}{\partial t}\right)_{\text{collisions}} = 0$$

In the first two terms, f may be replaced by the undisturbed Fermi distribution f_0. Evaluating these two terms, and using $f_0 = \dfrac{1}{e^{(E-J)/kT}+1}$ we get:

$$(13) \qquad \frac{df_0}{dx} = -\frac{\partial f_0}{\partial E}\left(\frac{\partial J}{\partial T} + \frac{E-J}{T}\right)\frac{dT}{dx}$$

and hence

$$(14) \qquad \left(\frac{\partial f}{\partial t}\right)_{\text{coll.}} = \frac{\partial f_0}{\partial E}v_x\left[eF - \left(\frac{\partial J}{\partial T} + \frac{E-J}{T}\right)\frac{dT}{dx}\right].$$

This equation can be solved by putting

$$(15) \qquad f(\mathbf{v}) = f_0(\mathbf{v}) + g(\mathbf{v}) = f_0(\mathbf{v}) - \left(\frac{\partial f}{\partial t}\right)_{\text{coll.}}\tau(v).$$

In equation (15) $\tau(v)$, the time of relaxation, is a constant independent of the direction of $\mathbf{v}$ but depending on its magnitude. $\dfrac{1}{\tau(v)}$ is a measure of the number of collisions suffered by an electron and is to be calculated from the integral equation

---

[*] Sommerfeld, A. & Bethe, H. A.   Handb. d. Phys. 24(2): 532.

$$(16) \qquad \left(\frac{\partial f}{\partial t}\right)_{\text{coll.}} = \int [g(\mathbf{v}') - g(\mathbf{v})] \, P\,(\mathbf{v} \cdot \mathbf{v}') \, dS'$$

where $\mathbf{v}'$ is the velocity of an electron with the *same* energy as $\mathbf{v}$, $P\,(\mathbf{v} \cdot \mathbf{v}')\,dS$ is the transition probability from $\mathbf{v}$ to $\mathbf{v}'$ per unit time and dS' is the element of solid angle associated with $\mathbf{v}'$. Substituting the value of $g(\mathbf{v})$ into the right-hand side of equation (16) and letting the left-hand side take the value given by equation (14) we get:

$$(17) \qquad \frac{1}{\tau(v)} = -\frac{1}{v_x} \int (v_x' - v_x)\, P\,(\mathbf{v} \cdot \mathbf{v}')\, dS'$$

Hence if we know $P\,(\mathbf{v} \cdot \mathbf{v}')$ we can find $\tau(v)$ and therefore $g(\mathbf{v})$ and $f(\mathbf{v})$. Thus

$$(18) \quad f(\mathbf{v}) = f_0(\mathbf{v}) - \tau(v) v_x \frac{\partial f_0}{\partial E}\left[eF - \left(\frac{\partial J}{\partial T} + \frac{E-J}{T}\right)\frac{dT}{dx}\right].$$

We can now insert this value of $f(\mathbf{v})$ into the equations (10) and (11) and find for j and Q, since the term in f_0 does not contribute, the following expressions:

$$(19) \quad j = \frac{2m^3}{h^3} \int v_x^2 \tau(v) \frac{\partial f_0}{\partial E}\left[eF - \left(\frac{\partial J}{\partial T} + \frac{E-J}{T}\right)\frac{dT}{dx}\right]d\mathbf{v}$$

$$(20) \quad Q = \frac{2m^3}{h^3} \int E v_x^2 \tau(v) \frac{\partial f_0}{\partial E}\left[eF - \left(\frac{\partial J}{\partial T} + \frac{E-J}{T}\right)\frac{dT}{dx}\right]d\mathbf{v}.$$

In the case of electron conduction in the stars, we may assume that the collision processes arise from the elastic scattering of the electrons by the atomic nuclei. If we let Θ be the angle between the velocity vectors $\mathbf{v}$ and $\mathbf{v}'$,

$$(21) \qquad P\,(\mathbf{v} \cdot \mathbf{v}') = P\,(\Theta, v) = N_A\, v\sigma\,(\Theta, v)$$

where $\sigma\,(\Theta, v) = + e^4 Z^2/m^2 v^4\,(1 - \cos\Theta)^2$ is the Rutherford differential cross-section and N_A the number of atoms per cubic centimeter. Inserting this expression we find:

$$(22) \qquad \frac{1}{\tau(v)} = -\frac{N_A e^4 Z^2}{m^2 v^3} \int\left(\frac{v_x'}{v_x} - 1\right) \cdot \frac{1}{(1 - \cos\Theta)^2}\, dS'.$$

Now we denote by ω the angle between $\mathbf{v}$ and the x-axis, by Θ, as before, the angle between $\mathbf{v}$ and $\mathbf{v}'$, and by ϕ the angle between the plane $\mathbf{vv}'$ and the plane containing $\mathbf{v}$ and the x-axis, so that:

$$v_x = v \cos\omega \qquad\qquad v_x' = v\,(\cos\omega\cos\Theta + \sin\omega\sin\Theta\cos\phi)$$

$$dS' = \sin \Theta \, d\Theta \, d\phi$$

and hence:

(23)
$$\frac{1}{\tau(v)} = -\frac{4\pi N_A e^4 Z^2 I}{m^2 v^3}$$

where we let
$$2I = \int \frac{\sin \Theta \, d\Theta}{(1 - \cos \Theta)}.$$

Putting $v_x = v \cos \omega$, $d\mathbf{v} = 2\pi v^2 dv \sin \omega \, d\omega$ and inserting the value (23) for $\tau(v)$ and the distribution function (15), (14) into (10), (11), we obtain

(24)
$$j = \int E^3 \left(\frac{\partial f_0}{\partial E}\right)\left[eF' - \frac{(E - J)}{T}\frac{dT}{dx}\right] dE = 0$$

(25)
$$Q = C \int E^4 \left(\frac{\partial f_0}{\partial E}\right)\left[eF' - \frac{(E - J)}{T}\frac{dT}{dx}\right] dE.$$

In (24) and (25) we have set:

(26)
$$eF' = eF - \frac{\partial J}{\partial T} \cdot \frac{dT}{dx}$$

and

(27)
$$C = \frac{16m}{3h^3 e^4 Z^2 N_A I}.$$

In order to derive the "second" approximation for the thermal conductivity in the degenerate case we must take $f_0 = \dfrac{1}{e^{(E-J)/kT}+1}$ and expand E^4 and E^3 around the point $E = J$. Thus since

$$E^4 = J^4 + 4J^3(E - J) + 6J^2(E - J)^2 + \ldots.$$

$$E^3 = J^3 + 3J^2(E - J) + 3J (E - J)^2 + \ldots.$$

we find on inserting (24) and (25) that:

$$\int_{-\infty}^{\infty} \frac{df_0}{d\varepsilon}\left(eF' - \varepsilon k \frac{dT}{dx}\right)[J^3 + 3J^2kT\varepsilon + 3Jk^2T^2\varepsilon^2 + k^3T^3\varepsilon^3]\, d\varepsilon = 0$$

or:

(28)
$$-J^3 eF' + \pi^2 k^2 T^2 eF'J - \pi^2 J^2 k^2 T \frac{dT}{dx} - \frac{7\pi^4}{15} k^4 T^3 \frac{dT}{dx} = 0$$

and

$$Q = C \int_{-\infty}^{\infty} \frac{df_0}{d\varepsilon}\left(eF' - k\varepsilon \frac{dT}{dx}\right)[J^4 + 4J^3kT\varepsilon + 6J^2k^2T^2\varepsilon^2 +$$
$$4Jk^3T^3\varepsilon^3 + k^4T^4\varepsilon^4]\, d\varepsilon$$

or

$$(29) \qquad Q = C \left\{ - J^4 eF' + 2\pi^2 J^2 k^2 T^2 eF' + \frac{7\pi^4}{15} k^4 T^4 eF' - \frac{4\pi^2}{3} k^2 T J^3 \right.$$

$$\left. \frac{dT}{dx} - \frac{28}{15} \pi^4 J k^4 T^3 \frac{dT}{dx} \right\}$$

In (28) and (29) we have taken $\varepsilon = \dfrac{E - J}{kT}$ and made use of the fact that:

$$\frac{df_0}{d\varepsilon} = - \frac{e^\varepsilon}{(1 + e^\varepsilon)^2} \qquad\qquad \text{and}$$

$$\int_{-\infty}^{\infty} \frac{e^z z^P dz}{(1 + e^z)^2} = 2\Gamma(P + 1) \cdot [1 - 2^{(1-P)}]\xi_P$$

<div align="right">when P is even
$= 0$ when P is odd.</div>

$\Gamma(P + 1) = P!$ and ξ_P is the Riemann zeta-function of argument P. Multiplying equation (28) by CJ and subtracting from equation (29) we get:

$$Q = C \left\{ \left(\pi^2 J^2 k^2 T^2 + \frac{7\pi^4}{15} k^4 T^4 \right) eF' - \left(\frac{\pi^2}{3} J^3 k^2 T + \frac{7\pi^4}{5} J k^4 T^3 \right) \frac{dT}{dx} \right\}$$

But since

$$eF' = - \frac{dT}{dx} \cdot \frac{\left(\pi^2 J^2 k^2 T + \dfrac{7\pi^4}{15} k^4 T^3 \right)}{(J^3 - \pi^2 J k^2 T^2)}$$

we get:

$$(30)\ Q = - \frac{C\pi^2 k^2 T^2}{3TJ} \left[\frac{\left(J^6 + \dfrac{31}{5} \pi^2 k^2 T^2 J^4 - \dfrac{7}{5} \pi^4 k^4 T^4 J^2 + \dfrac{49}{75} \pi^6 k^6 T^6 \right)}{J^2 - \pi^2 k^2 T^2} \right] \frac{dT}{dx}$$

or

$$(31)\ \lambda = \frac{16\pi^2 m k^2 T}{9h^3 e^4 Z^2 N_A I J} \left[\frac{\left(J^6 + \dfrac{31}{5} \pi^2 k^2 T^2 J^4 - \dfrac{7}{5} \pi^4 k^4 T^4 J^2 + \dfrac{49}{75} \pi^6 k^6 T^6 \right)}{(J^2 - \pi^2 k^2 T^2)} \right]$$

We wish to know λ to terms of the order $\left(\dfrac{kT}{J_0} \right)^2$ where J_0 is the Fermi point-energy at absolute zero. Making use of the relation $J = J_0 - \dfrac{\pi^2 k^2 T^2}{12 J_0} \cdots$ we find as a second approximation to λ

$$(32) \qquad \lambda^{D.} = \frac{16\pi^2 m J_0^3 k^2 T}{9h^3 e^4 Z^2 N_A I} \left(1 + 6.95\ \pi^2 \frac{k^2 T^2}{J_0^2} \right);$$

By equation (9) the degenerate conductive opacity is now determined. If we insert numerical values for the physical constants and the value 6 for $\left(\dfrac{Z^2}{A}\right)$ ("Russell" mixture), then we find:

$$(33) \qquad K_c^{\text{D.}} = \frac{7.82 \cdot 10^{-7} X_R I T^2}{(\rho\eta_e)^2(1 + 7.74 \cdot 10^{-10} W^2)}$$

we have written $J_0/kT = \dfrac{2.98 \cdot 10^5}{W}$ where $W = \dfrac{T}{(\rho\eta_e)^{2/3}}$ and η_e is related to the free electron density N_e, viz. $N_e = \dfrac{\rho\eta_e}{m_H}$. Further we put

$N_A = \dfrac{\rho}{m_H A}$ where A is the atomic weight. The above expression for the degenerate conductive opacity is suitable for application.

We wish now to derive an expression for the non-degenerate conductive opacity. In this case $f_0 = \gamma e^{-E/kT} = \gamma e^{-\varepsilon}$ (γ is really $e^{-\Psi/kT}$) where $\varepsilon = E/kT$. We now consider equations (24) and (25); we have

$$(34) \qquad j \approx \int_0^\infty \varepsilon^3 \frac{\partial f_0}{\partial \varepsilon}\left(eF'' - k\varepsilon\frac{dT}{dx}\right) d\varepsilon = 0$$

$$(35) \qquad Q = C\,(kT)^4 \int_0^\infty \varepsilon^4 \frac{\partial f_0}{\partial \varepsilon}\left(eF'' - k\varepsilon\frac{dT}{dx}\right) d\varepsilon$$

with

$$eF'' = eF' + J/T\frac{dT}{dx}$$

equation (34) gives:

$$\int_0^\infty \varepsilon^3 e^{-\varepsilon}\left(eF'' - k\varepsilon\frac{dT}{dx}\right) d\varepsilon = \left[3!eF'' - 4!k\frac{dT}{dx}\right] = 0$$

or

$$eF'' = 4k\frac{dT}{dx}$$

when this is used in conjunction with equation (35) we find:

$$(36) \qquad Q = -\,4!C\gamma k^5 T^4 \frac{dT}{dx}.$$

Since C is given by (27) and γ is only $e^{-\Psi/kT}$, we obtain:

$$(37) \qquad \lambda^{\text{N.D.}} = \frac{128 m k^5 T^4 e^{-\Psi/kT}}{h^3 e^4 Z^2 N_A I}$$

again, noting that $e^{-\Psi/kT} = 8.25 \cdot 10^{-9} W^{3/2} - .354$ and inserting numerical values for constants, we get for the non-degenerate conductive capacity:

$$(38) \qquad K_c\text{N D.} = \frac{2.85 \cdot 10^{+9} X_R I}{T} \cdot (8.25 \cdot 10^{-9} W^{3/2} - .354)$$

Thus far we have not given any explicit form for I. If we look at the integral expression for I (cf. equation (23)) we see that it diverges if we integrate between the limits 0 and π. However, there are physical grounds for extending this integration only to some small angle Θ_0, in which case:

$$(39) \qquad 2I = \log_e \frac{2}{(1 - \cos \Theta_0)}.$$

Now it can be shown that $\Theta_0 \approx \dfrac{\lambdabar}{a}$ * where $\lambdabar$ is the de Broglie wavelength of the electron participating in the collision, and "a" is the screening radius of the atom. A reasonable value for "a" is given by taking $\dfrac{4\pi a^3}{3} = \dfrac{Z}{2N_e}$ (Z/N_e is the approximate atomic volume associated with the charge Z). Since $\lambdabar = \dfrac{h}{2\pi m v}$ we get:

$$(40) \qquad I = \log_e \left[\frac{4\pi m v}{h} \cdot \left(\frac{3Z}{8\pi N_e} \right)^{1/3} \right]$$

In the degenerate case

$$\frac{m v^2}{2} = J_0 - \frac{\pi^2 k^2 T^2}{12 J_0} \qquad N_e = W^{3/2} m_H$$

$$J_0/kT = \frac{2.98 \cdot 10^5}{W},$$

so that we have

$$(41) \qquad I\text{D.} = \frac{1}{3}\left[5.9 + \frac{3}{2}\log_e (1 - 9.3 \cdot 10^{-12} W^2) \right]$$

In the non-degenerate case

$$\frac{m v^2}{2} = \frac{3}{2} kT \left(1 + \frac{2.24 \cdot 10^7}{W^{3/2}} \right)$$

and we find for I:

$$(42) \quad I\text{N.D.} = \frac{1}{3}\left[\log_e (3.92 \cdot 10^{-6} W^{3/2}) + \log_e \left(1 + \frac{2.15 \cdot 10^7}{W^{3/2}} \right) \right]$$

These expressions are to be taken in conjunction with (33) and (38).

* **Bethe, H. A.** Handbuch d. Phys. 24(1): 497.

The above results have been used to integrate the star equations for the white dwarfs, Sirius B and 40 Eridani B, with some very interesting consequences regarding the energy-production process and the radii of these stars. These further investigations will be found in a paper in the Astrophysical Journal.[†]

The author wishes to thank Professor H. A. Bethe for helpful discussion.

[†] Marshak, R. E. Ap J. **92:** 321 (1940); in this paper will also be found "third" approximations for the conductive opacity

ON STELLAR ENVELOPES

By Jaakko Tuominen*

From the Harvard College Observatory, Cambridge, Massachusetts

INTRODUCTION

The rate of liberation of subatomic energy, which follows from the recent theories of nuclear processes in the stars, is very sensitive to the temperature.[†] Therefore the knowledge of the internal temperature plays a very important rôle in physical studies of the stellar interior. The calculation of the internal temperatures of observed stars is itself connected with the knowledge of the absorption coefficient in the outer layers. This relationship can be explained in the following way: First, the central temperature depends on the mass concentration in the star, so that a higher mass concentration demands a higher central temperature. Further, the mass concentration is determined by the extension of the outer layers of a star, which have a low density and contribute only a small fraction of the stellar mass, but a large fraction of the volume and radius. When the absorption coefficient in this envelope is high, the envelope must necessarily be more extended than for a lower coefficient of absorption.

The central temperatures generally used in physical studies of energy generation in the stars are based on the standard model. The right to use this model, as far as the main-series stars are concerned, is strongly confirmed by recent studies of the mass concentration in eclipsing binaries.[‡] Furthermore, when the main-series stars are assumed to have the standard model constitution, Bethe's theory[§] of energy generation provides an adequate explanation for the energy which these stars radiate. The standard-model temperatures of giants, however, are so low that Bethe's theory cannot explain their high luminosity; and although attempts have been made[¶] to explain the energy generation in giants with these low temperatures, the process still seems to be obscure. The problem might be simplified if we could

---

show that under certain conditions very large stellar envelopes and consequently high central temperatures are physically possible.

For low densities the Kramers-Eddington law of opacity appears to be well justified, at least when the chemical composition is supposed to be constant throughout the envelope. It is interesting, however, to speculate upon the different kinds of envelopes we shall find when the absorption coefficient has the form

$$\kappa = \kappa_1 \rho^p / T^q \tag{1}$$

where ρ and T are density and temperature, and κ_1, p, and q are constants. In our problem, p and q will be considered as parameters.

FUNDAMENTAL EQUATIONS

We assume, further, that the radiation field is nearly isotropic. This assumption is correct if the optical depth be great enough. Further, it is correct to assume that the envelope does not contain energy sources and that the mass of the envelope is negligible in comparison to the stellar mass. Then the following equations* determine the constitution of the envelope:

$$\frac{d}{dr}\left(\frac{R}{\mu}\rho T + \frac{a}{3}T^4\right) = -\frac{GM\rho}{r^2}; \tag{2}$$

$$\frac{16\pi a c T^3}{3\kappa\rho}\frac{dT}{dr} = -\frac{L}{r^2} \tag{3}$$

Here, r is the radius vector from the stellar center, M and L the mass and luminosity of the star, and μ the mean molecular weight. The latter quantity has been assumed constant throughout the envelope. R is the gas constant, a the density of radiation at $1°$ K, and c the velocity of light. $\frac{R}{\mu}\rho T$ represents the gas pressure and $\frac{a}{3}T^4$ the radiation pressure. Equation (2) is the condition for hydrostatic equilibrium and equation (3) the equation of radiative transfer.

For further analysis, we eliminate r from (2) and (3) and substitute κ as given by equation (1). The following equation is then obtained:

$$\frac{d(\rho T)}{dT} = \left(\frac{16\pi a c G\mu M}{3R\kappa_1 L}\frac{T^q}{\rho^p} - \frac{4a\mu}{3R}\right)T^3, \tag{4}$$

which gives the relation between density and temperature in the stellar envelope. The value of the radius vector corresponding to a certain density or temperature is determined by relation (3) or

* Eddington, "The Internal Constitution of the Stars." Cambridge. 1930. p. 114.

$$\frac{dT}{dr} = -\frac{3\kappa_1 L}{16\pi ac}\frac{1}{r^2}\frac{\rho^{p+1}}{T^{q+3}}.$$ (5)

LIMITING VALUES OF q/p

Here we shall attempt to show* that stellar envelopes are essentially different according as $q/p < 3$, $q/p = 3$, $3 < q/p < 4$, or $q/p \geqq 4$.

(1) When $q/p < 3$, the temperature gradient $\dfrac{dT}{dr}$ at the surface—that is, where the temperature vanishes—is finite. If the values assigned to q and p change, the change in the distribution of temperature and density in the envelope is of the same order of magnitude as the change in q/p. In the part of the envelope nearest to the surface, the radiation pressure may be neglected in comparison to the gas pressure.

(2) When $q/p = 3$, the conditions in the envelope are similar to those in the case $q/p < 3$, except that the ratio between radiation pressure and gas pressure is finite at the surface.

(3) When $3 < q/p < 4$, $\dfrac{dT}{dr}$ at the surface vanishes. When the value of q/p increases, the volume occupied by the envelope increases relatively fast. Hence also the mass concentration of the star increases relatively fast. The gas pressure is negligible in comparison to radiation pressure at the surface.

(4) When q/p approaches 4, the stellar radius, as well as the mass concentration, approach infinity. Also, the total mass of the envelope increases towards infinity. For all models with $q/p \geqq 4$, the radius, mass, as well as the mass concentration, are infinitely large. The gas pressure is negligible at sufficiently large distances from the stellar center.

These results are derived in the following way:

First, the density is developed in series according to powers of the temperature,

$$\rho = \alpha T^\nu + \ldots \ldots$$ (6)

where α and ν are positive numbers, and ν is the lowest power of T in the series. By substitution of (6), equation (4) assumes the form

$$(\nu + 1)\, \alpha T^{\nu-3} + \ldots = -\frac{4a\mu}{3R} + \frac{16\pi acG\mu M}{3R\kappa_1 L\alpha^p}\, T^{q-\nu p} + \ldots . \quad (7)$$

On the left-hand side of this equation, the term written down is the term of lowest degree. On the right-hand side the conditions are more complicated, because either the first or the second term, or their sum, may represent the term of lowest degree. We have then three different cases according as $\nu \lesseqgtr 3$. As will be shown, the sign $<$ and the sign of equality correspond to $q/p < 3$ and $q/p = 3$ (page 63), respectively, while the lower sign corresponds to $q/p > 3$.

(a) When $\nu < 3$, equation (7) demands that

$$(\nu + 1)\alpha T^{\nu-3} = \frac{16\pi acG\mu M}{3R\kappa_1 L\alpha^p}\, T^{q-\nu p}.$$

Hence

$$\nu = \frac{q + 3}{p + 1}, \quad (8)$$

and the constant α is determined by the relation

$$\alpha^{p+1} = \frac{16\pi acG\mu M}{3R\kappa_1 L(\nu + 1)}.$$

Formula (8) shows that this case occurs when

$$q/p < 3.$$

Formulae (6) and (8) now give the dependence of ρ on T:

$$\rho = \alpha T^{\frac{q+3}{p+1}} + \ldots ; \quad (9)$$

and this formula, together with (5), yields the temperature gradient

$$\frac{dT}{dr} = -\frac{3\kappa_1 L}{16\pi ac}\frac{1}{r^2}\,(\alpha^{p+1} + \ldots), \quad (10)$$

where the omitted terms contain positive powers of T. The latter relation shows that the temperature gradient at the stellar surface is finite and, further, that a small change in p or q produces a small change also in $\dfrac{dT}{dr}$. This result was stated on page 63.

The relation between the radius vector r and the temperature T is found by integration of equation (10). We write it first in the form

$$(1 + \ldots)dT = -\frac{3\kappa_1 L\alpha^{p+1}}{16\pi ac}\frac{dr}{r^2},$$

where the omitted terms on the left-hand side contain positive powers of T. As we assume that the temperature vanishes at the surface, where the radius vector is R, it follows that

$$\frac{1}{r} - \frac{1}{R} = \frac{16\pi ac}{3\kappa_1 L\alpha^{p+1}}\, T + \cdots,$$

in which the omitted terms on the right-hand side contain higher powers of T.

The ratio of radiation pressure to gas pressure is

$$\frac{a\mu}{3R}\, \frac{T^3}{\rho}.$$

As may be seen from formula (9), in the envelope this quantity is proportional to

$$T^{\frac{3p-q}{p+1}} + \cdots,$$

which shows that the radiation pressure vanishes at the stellar surface.

(b) When $\nu = 3$, all terms written down in formula (7) are independent of T. The equation demands that

$$q/p = 3,$$

and α is determined by the equation

$$(\nu + 1)\alpha = -\frac{4a\mu}{3R} + \frac{16\pi acG\mu M}{3R\kappa_1 L\alpha^p}.$$

In other respects, however, the series determining ρ and dT/dr as functions of T, are, as far as the first terms are concerned, exactly similar to the corresponding formulae in the case $\nu < 3$. Therefore, when $q/p = 3$, the temperature gradient at the stellar surface is likewise finite. The ratio between radiation pressure and gas pressure is finite at the surface, since in this case

$$\rho \propto T^3 + \cdots.$$

(c) and (d) When $\nu > 3$, it is evident from equation (7) that necessarily

$$\nu = q/p. \tag{11}$$

Hence this case occurs when

$$q/p > 3.$$

Further, α is determined by the relation

$$\alpha^p = \frac{4\pi cGM}{\kappa_1 L}.$$

(12)

Formulae (6) and (11) give the dependence of density on temperature:

$$\rho = \alpha \, T^{q/p} + \ldots$$

(13)

Hence it follows, from relation (5), that

$$\frac{dT}{dr} = -\frac{3\kappa_1 L}{16\pi ac} \frac{1}{r^2} (\alpha^{p+1} T^{q/p-3} + \ldots),$$

(14)

where, as in the earlier series, the omitted terms contain higher powers of T. This relation shows that the temperature gradient vanishes at the stellar surface, and that a small change in q or p is sufficient to change the constitution of the envelope significantly. When q/p increases, the mass concentration also increases, as was stated on page 63. Further, the ratio between radiation pressure and gas pressure is proportional to

$$T^{3-q/p} + \ldots ,$$

and hence the gas pressure is negligible in comparison to radiation pressure in the part of the envelope nearest to the stellar surface.

Relation (14) is now written in the form

$$(T^{-(q/p-3)} + \ldots) \, dT = -\frac{3\kappa_1 L \alpha^{p+1}}{16\pi ac} \frac{dr}{r^2}.$$

(15)

From this equation we see immediately that the conditions in the envelope are very different according as $q/p < 4$ or $q/p \geqq 4$. We consider separately these two different cases, first mentioned on page 63.

For $q/p < 4$, (c), equation (15) may be integrated, and the radius vector is determined as a function of the temperature by the equation

$$\frac{1}{r} - \frac{1}{R} = \frac{1}{4 - q/p} \frac{16\pi ac}{3\kappa_1 L \alpha^{p+1}} T^{4-q/p} + \ldots ,$$

where R is the stellar radius.

For $q/p \geqq 4$, (d), however, the integral of the left-hand side of formula (15) approaches infinity when the lower limit of integration—that is, the surface temperature—approaches zero. But the integral of the right-hand side is finite if $L \, \alpha^{p+1}$ is finite. Using formula (12), we obtain

$$L\alpha^{p+1} = \left(\frac{4\pi cG}{\kappa_1}\right)^{1+1/p} \frac{M^{1+1/p}}{L^{1/p}}.$$

Hence, in order that both sides of equation (15), when integrated from $T = 0$, may be equal, M must necessarily approach infinity, when T approaches zero; from which it follows that, in stellar models where $q/p \geq 4$, the envelopes are extended to infinity; and the masses of the envelopes, and hence also the masses of these stellar models, are infinitely large.

Even without mathematical analysis it is easy to understand why conditions like those derived in this section can exist. We may illustrate it in the following way:

First we assume that the absorption coefficient

$$\kappa = \kappa_1 \rho^p / T^q$$

is so small for small values of the temperature and density that the force of the radiation pressure (which is proportional to the absorption coefficient) is negligible in comparison with the force of the gas pressure. Evidently this condition is possible when q is small in comparison to p. If, then, q increases or p decreases, the value of κ, for vanishing density and temperature, increases. Hence the importance of the radiation pressure increases. When q increases (or p decreases) still more, the values of the absorption coefficient may become so large that the gas pressure may be neglected in comparison with the radiation pressure. The detailed theory showed that this condition arises when $q/p > 3$; radiation pressure is negligible in comparison with gas pressure when $q/p < 3$, while both pressures are of the same order of magnitude when $q/p = 3$. Let us now assume that q increases (or p decreases) still further. Then the radiation pressure may possibly become so strong that it can be counterbalanced only by an infinitely large mass. According to the exact theory, this happens when $q/p \geq 4$.

FORMULAE FOR NUMERICAL CALCULATION OF THE CONSTITUTION OF THE ENVELOPE

The radii of convergence of the series developed in the above are too small to facilitate a numerical calculation of the constitution of the envelope. For this purpose we propose to derive in this section more accurate formulae which will represent the density, temperature, and radius as functions of a certain parameter. A procedure similar to that used by Chandrasekhar[*] and by the present author[†] will be fol-

[*] Chandrasekhar, M. N. 96: 647. 1936.

[†] Ann. Acad. Scient. Fennicae, A 48: No. 16. 61–64. 1938.

lowed. In order to make the problem somewhat simpler, we choose for p the value unity.

We introduce in equation (4) a new function ϑ, defined by the equation

$$\rho = \frac{a\mu}{3R} \vartheta T^3. \tag{16}$$

Equation (4) then takes the form

$$\frac{1}{4} \vartheta T \frac{d\vartheta}{dT} = \frac{4\pi cGM}{\kappa_1 L} \frac{3R}{a\mu} T^{q-3} - \vartheta (1 + \vartheta).$$

Substituting

$$z = \frac{4\pi cGM}{\kappa_1 L} \frac{3R}{a\mu} T^{q-3}, \tag{17}$$

we obtain the above relation in the form

$$\frac{1}{4} (q - 3) \vartheta z \frac{d\vartheta}{dz} = z - \vartheta (1 + \vartheta);$$

and, further, when

$$\frac{1}{4} (q - 3) = e,$$

the following differential equation results:

$$e\vartheta z \frac{d\vartheta}{dz} = z - \vartheta (1 + \vartheta). \tag{18}$$

As in our problem we consider only such q values as are larger than 3 but smaller than 4, e is a small quantity and we may write:

$$\vartheta = \vartheta_0 + e\vartheta_1 + e^2 \vartheta_2 + \ldots \ldots \tag{19}$$

Inserting this in (18) and equating the powers of e, we obtain

$$z = \vartheta_0 (1 + \vartheta_0); \tag{20}$$

$$- z\vartheta_0 \frac{d\vartheta_0}{dz} = (1 + \vartheta_0) \vartheta_1 + \vartheta_0\vartheta_1; \tag{21}$$

Combining (19), (20), and (21), we have the following series:

$$\vartheta = \vartheta_0 \left(1 - \frac{e}{4} \frac{z}{z + \frac{1}{4}} + \ldots \right).$$

As $\dfrac{e}{4} = \dfrac{q-3}{16}$, the second term does not, even in the least favorable case $(q = 4)$, contribute more than 6.25 per cent. Consequently, in further work the approximation

$$\vartheta = \vartheta_0$$

will be used. According to (16), therefore,

$$\rho = \frac{a\mu}{3R}\,\vartheta_0 T^3. \tag{22}$$

Equation (17) and (20), when combined, give

$$T = \left[\frac{a}{3R}\,\frac{L}{4\pi cGM}\,\mu\kappa_1\vartheta_0\,(\vartheta_0 + 1)\right]^{\frac{1}{q-3}}. \tag{23}$$

Hence

$$\rho = \left(\frac{a}{3R}\right)^{\frac{q}{q-3}}\left(\frac{L}{4\pi cGM}\right)^{\frac{3}{q-3}}\mu^{\frac{q}{q-3}}\kappa_1^{\frac{3}{q-3}}\vartheta_0^{\frac{q}{q-3}}\,(\vartheta_0 + 1)^{\frac{3}{q-3}}. \tag{24}$$

The last two formulae give the temperature and density as functions of the parameter ϑ_0. In order to find the corresponding values of the radius vector, differential equation (5) must be integrated. We introduce ρ and T as functions of ϑ_0, as given by (22) and (23).

Then equation (5) assumes the form

$$\frac{dr}{r^2} = -\frac{16\pi ac}{3L}\left(\frac{a}{3R}\right)^{\frac{4-q}{q-3}}\left(\frac{L}{4\pi cGM}\right)^{\frac{q-2}{q-3}}\mu^{\frac{4-q}{q-3}}\kappa_1^{\frac{1}{q-3}}\frac{1}{q-3}$$
$$\times\,\vartheta_0^{\frac{7-2q}{q-3}}\,(\vartheta_0 + 1)^{\frac{1}{q-3}}\,(2\vartheta_0 + 1)\,d\vartheta_0 \tag{25}$$

In order to see what the value of ϑ_0 is at the stellar surface, we consider equations (22) and (13). The latter shows that in the vicinity of the surface $\rho\alpha T^q$. Hence, according to (22), ϑ_0 approaches zero at the surface. Therefore it follows from formula (25) that

$$\frac{R}{r} - 1 = \frac{16\pi acR}{3L}\left(\frac{a}{3R}\right)^{\frac{4-q}{q-3}}\left(\frac{L}{4\pi cGM}\right)^{\frac{q-2}{q-3}}\mu^{\frac{4-q}{q-3}}\kappa_1^{\frac{1}{q-3}}\frac{1}{q-3}$$
$$\times\int_0^{\vartheta_0}\vartheta_0^{\frac{7-2q}{q-3}}\,(\vartheta_0 + 1)^{\frac{1}{q-3}}\,(2\vartheta_0 + 1)\,d\vartheta_0. \tag{26}$$

Formulae (23), (24) and (26) will, in the following, be used for the study of the constitution of the stellar envelopes.

Generally, the integral in equation (26) must be calculated numerically. In the case of the Kramers-Eddington law of opacity ($q = 3.5$), however, its analytical evaluation is very simple. Thus[*]:

$$\frac{R}{r} - 1 = \frac{16\pi a c R}{3L}\left(\frac{a}{3R}\right)\left(\frac{L}{4\pi c G M}\right)^3 \mu \kappa_1^2 \cdot \frac{1}{3}[(\vartheta_0+1)^3(3\vartheta_0+1) - 1] \quad (27)$$

while

$$T = \left[\frac{a}{3R}\frac{L}{4\pi c G M}\mu\kappa_1\vartheta_0\,(\vartheta_0 + 1)\right]^2 \quad (28)$$

and

$$\rho = \left(\frac{a}{3R}\right)^7\left(\frac{L}{4\pi c G M}\right)^6 \mu^7 \kappa_1^6 \vartheta_0^7\,(\vartheta_0 + 1)^6. \quad (29)$$

THE CONSTANT FACTOR IN THE COEFFICIENT OF OPACITY

The constant factor κ_1 (formula (1)) was determined so that the coefficient of opacity (for $p = 1$)

$$\kappa = \kappa_1\frac{\rho}{T^q}, \quad (30)$$

at a certain temperature and density, is equal to the coefficient of opacity found by physical calculations. P. M. Morse's[†] calculations reach to temperatures as low as 400,000 degrees and to densities as low as $\rho = 0.01/(1 - X)$, where X is the hydrogen abundance. These calculations show that for the Russell mixture

$$\kappa = [25.64]\frac{1 - X^2}{\tau}\frac{\rho}{T^{3.5}}, \quad (31)$$

where τ, the guillotine factor, has been tabulated by Morse for different densities and temperatures. In order to make expressions (30) and (31) equal, we must have

$$\kappa_1 = [25.64]\frac{1 - X^2}{\tau}T^{q-3.5}. \quad (32)$$

Morse's calculations show that for $\rho = 0.01/(1 - X)$, τ is at a minimum at about $T = 500,000°$. At this temperature $\tau = 1.25$. When these values are substituted in (32), it is found that

$$\kappa_1 = [25.54]\,(1 - X^2)\,(500000)^{q-3.5}. \quad (33)$$

Further, it seems that the most likely value for the mean molecular

weight in the stars[*] is $\mu = 1$. μ and X are related to each other by the equation:[†]

$$X = \frac{\frac{1}{\mu} - 0.5}{1.5}. \tag{34}$$

Hence the X-value corresponding to $\mu = 1$ is $X = 0.33$. Substituting this in (33) we find

$$\kappa_1 = [25.5] \ (500000)^{q-3.5}. \tag{35}$$

This expression for κ_1 will be applied in the following calculations.

THE ENVELOPE OF CAPELLA A

The above theory was applied to the brighter component of α Aurigae, one of the best observed giants. The logarithms of its mass, luminosity, and radius, in units of the solar mass, are:[‡]

$$\log (M/M_\odot) = 0.62;$$

$$\log (L/L_\odot) = 2.08;$$

$$\log (R/R_\odot) = 1.20.$$

When we assume that the distribution of density and temperature in this star is the same as that in the standard model, the following central temperature is found:

$$T_c = 5.1 \text{ million degrees.}$$

The corresponding value for the sun is $(T_c)_\odot = 19.5$ million degrees.

Using the above values for the mass, luminosity, and radius, and assuming different values for the quantity q, the author has calculated the constitution of the envelope by means of formulae (23), (24), and (26), and, in the case $q = 3.5$, by means of formulae (27), (28), and (29). It was assumed that a third of the mass of the envelope is hydrogen and, further, that at the density[§] $\rho = 0.015$ gr/cm[3] and temperature 500000° K, the absorption coefficient is equal to the absorption coefficient for this density and temperature as given by Morse's calculations for Russell-mixture (compare preceding section). Then the

---

[*] Strömgren, B. Z. f. Ap. 7: 235 (Table 5). 1933.
[†] Strömgren, B. Z. f. Ap. 7: 229. 1933.
[‡] Chandrasekhar, "An Introduction to the Study of Stellar Structure" Chicago, Ill. p. 489. 1939.
[§] If Capella A is assumed to have the standard-model distribution of density, its central density is 0.080 gr/cm[3].

coefficient κ_1 for each q-value, respectively, may be calculated from formula (35).

The result of the calculations is shown in FIGURE 1. The stellar center is at the right while the surface is at the left. All the curves approach the same line, $r/R = 1$, asymptotically.

The diagram clearly shows how the mass concentration increases when q approaches 4.* There is, however, an effect which to some extent counterbalances this increase of the mass concentration with increasing q. In the diagram, as in the above calculations, the surface is the level where density and temperature vanish. However, it seems more correct to define the stellar surface as the farthermost level to which the sight of an observer outside the star can penetrate. The radius of this spherical surface may be defined as a solution, r, of the equation:

$$\int_{R}^{r} \rho\kappa dr = 1, \tag{36}$$

where all letters have the same meanings as before. When q increases, the solutions r of this equation decrease because the values of ρ for different r-values decrease.
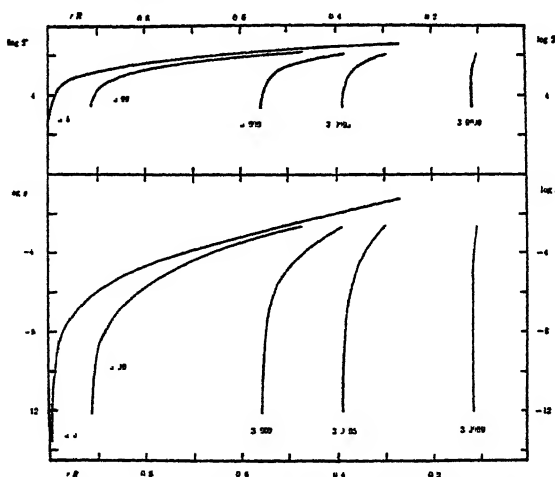
FIGURE 1. The distribution of temperature and density in the envelope of Capella A for different values of q ($\kappa = \kappa_1 \rho/T^q$).

* That a change of this kind in the law of opacity might make the central temperatures of giants higher than their standard-model values was suggested to the author by Professor S. Rosseland.

THE MASS CONCENTRATION AS A FUNCTION
OF THE STELLAR MASS

Under the discussion of limiting values of q/p it was shown that, for absorption coefficients where $3 < q/p < 4$, the gas pressure may be neglected in comparison to radiation pressure in the layers nearest to the surface. This result is of special interest as the Kramers-Eddington law of opacity belongs to this region. Although the gas pressure may be neglected actually in only a very small region, still it is interesting to consider models of stellar envelopes in which the sole pressure is the radiation pressure. Such a study showed (it was assumed that $p = 1$) that the mass concentration in heavy stars increases with increasing mass as well as with increasing radius. Account was taken of the observational fact that for heavy stars[*]

$$L \alpha \, M^{3\frac{1}{2}}.$$

The increase of the mass concentration with mass and radius can be made as large as we like by choosing a value for q sufficiently near to four. At the same time, as was shown above, the mass concentration itself becomes high. This increase of the mass concentration with stellar mass and radius Chandrasekhar[†] found from observational data by numerical calculations of the structure of the envelope of certain stars. He used the Kramers-Eddington law of opacity and his calculations show that the red giants are more centrally condensed than other stars.

FINITE SURFACE TEMPERATURE AND DENSITY

In the above investigations it was assumed that the temperature and density vanish at the stellar surface. Although this assumption evidently is not correct for actual stars, it is generally believed that introduction of correct surface temperatures and densities does not alter the internal constitution of a stellar model, because in any case they are very low in comparison to the densities and temperatures inside the stars.

The surface temperatures of many stars can be determined with considerable certainty from their spectra or colors. (By 'surface' is meant the photospheric layers of the stars.) The theories on which the determination of surface density is based are not so well confirmed. If we consider the surface temperature of a star to be a known constant

---

[*] Kuiper, Ap. J. **88**: 506. 1938.
[†] "Introduction to the Study of Stellar Structure." Chicago, Ill. 318–319. 1939.

and assume sufficiently low values for the density at the surface, formulae (2), (3), and (1) show clearly that both dT/dr and $d\rho/dr$ can be as small as we like in the outer parts of a star. Therefore, when we assume a sufficiently low value for the surface density, the star may have a very high mass concentration. A closer investigation shows, however, that such an envelope is for the most part so transparent that the actually observable surface of the star lies far inside the envelope.

The solutions of equations (2) and (3), where the temperature does not vanish, have the nature that the density approaches zero asymptotically, while the mass approaches infinity when the radius vector increases.* In spite of this fact, the mass in the envelope may be considered as independent of the radius vector, if very large distances from the stellar center are excluded. The author has tried to determine the level in such an envelope to which the sight of an outside observer can penetrate.

In order to simplify the problem, it will be assumed that the temperature is constant (T_{surf}) down to the observable surface. Then equation (2) assumes the form

$$\frac{R}{\mu} T_{surf} \frac{d\rho}{dr} = - GM \frac{\rho}{r^2}.$$

The radius R corresponds to ρ_{surf} while r corresponds to ρ. Therefore, the integration gives

$$\rho = \rho_{surf}\, e^{-\frac{G\mu M}{R T_{surf}}\left(\frac{1}{R} - \frac{1}{r}\right)}.$$

When this expression is substituted in the integral (36), it follows, when the Kramers-Eddington law of opacity ($\kappa = \kappa_1\, \rho/T^{3.5}$) is assumed to hold, that

$$\int_R^r \rho\kappa\, dr = \kappa_1 \frac{\rho_{surf}^2}{T_{surf}^{3.5}} e^{-\frac{2G\mu M}{R T_{surf} R}} \int_R^r e^{\frac{2G\mu M}{R T_{surf}} \frac{1}{r}}\, dr$$

In order to simplify the integration we make the substitution

$$\frac{2G\mu M}{R T_{surf}} \frac{1}{r} = X.$$

Then:

$$\int_R^r \rho\kappa\, dr = \kappa_1 \frac{\rho_{surf}^2}{T_{surf}^{3.5}} e^{-\frac{2G\mu M}{R T_{surf} R}} \frac{2G\mu M}{R T_{surf}} \int_{\frac{2G\mu M}{R T_{surf}} \frac{1}{r}}^{\frac{2G\mu M}{R T_{surf}} \frac{1}{R}} e^x \frac{1}{x^2}\, dx. \tag{37}$$

---

* Tuominen, Ann. Acad. Scient. Fennicae A 43: 18–21. 1938.

The numerical value of the expression

$$\frac{2G\mu M}{R T_{surf}} \frac{1}{R} \tag{38}$$

has been calculated for Capella A. As the masses, luminosities, and radii were taken form Chandrasekhar's table,[*] the effective temperatures were calculated by means of the formula[†]

$$L = \pi a c \, R^2 T_{eff}^{4}$$

The surface temperature is then given by the relation[‡]

$$T_{surf} = \frac{1}{\sqrt[4]{2}} T_{eff}.$$

For Capella A the calculation showed that log T_{eff} = 3.68, while log T_{surf} = 3.60; and the quantity (38) was found to be 3020. This means that, in formula (37), r need be only a very little greater than R, and the integrand at the lower limit will then practically vanish in comparison to its value at the upper limit. Between the limits of integration, e^x will change very fast while $1/x^2$ remains practically constant. Therefore, we can place $1/x^2$ in front of the integral sign and substitute for it the upper limit of integration. Then the integration can be carried out, and the lower limit may be neglected in comparison to the upper limit. The following result is found:

$$\int_{R}^{r} \rho \kappa \, dr = \frac{\rho_{surf}^2}{T_{surf}^{2.5}} \frac{\kappa_1 R}{2G\mu} \frac{R^2}{M}.$$

This quantity we write equal to unity. Then the radius R of the star will be defined as the radius of the farthermost level in the envelope to which sight can penetrate from the space around the star. Then it follows that

$$\rho_{surf}^2 = \frac{2G\mu}{\kappa_1 R} \frac{M}{R^2} T_{surf}^{2.5}. \tag{39}$$

The calculation of ρ_{surf} involves the knowledge of μ and κ_1. These quantities are, of course, different outside the star than inside. We shall, however, write $\mu = 1$, log $\kappa_1 = 25.5$, as we did in the calculations with vanishing surface temperature and density. The inaccuracy of the assumption should not be very serious as long as we only compare

---

[*] "An Introduction to the Study of Stellar Structure." Chicago, Ill. p. 489. 1939.
[†] Eddington "The Internal Constitution of the Stars." Cambridge. p. 120. 1930.
[‡] Eddington "The Internal Constitution of the Stars." Cambridge. p. 323. 1930.

stellar envelopes with finite surface density and temperature to those with vanishing surface temperature and density. Formula (39) gives for log $T_{surf} = 3.60$, log $\rho_{surf} = -10.73$. When we calculate the density corresponding to log $T = 3.60$ from formulae (28) and (29), we find log $\rho = -11.22$. This density, which occurs very near the stellar surface, is of the same order as ρ_{surf} calculated above, showing that the part of the envelope constructed by using finite surface temperatures instead of vanishing surface temperatures is so transparent that the effective radius is practically unchanged. Hence it follows that the mass concentration cannot be sensibly changed by taking into account the correct boundary condition instead of the boundary condition $\rho = 0,\; T = 0$.

SUMMARY

It has been shown that the mass concentrations and therefore the central temperatures of stars will be much higher than the standard-model values if the ratio between the exponent of the temperature and the exponent of the density in the opacity coefficient ($\kappa = \kappa_1 \rho^p / T^q$) is higher than 3.5, which is its value in the Kramers-Eddington law of opacity. If the ratio approaches 4, the ratio between the central density and the mean density of the star approaches infinity. Further, it has been shown that although the introduction of finite surface temperatures and densities may lead to extended envelopes, their densities are necessarily so low that the envelopes are transparent, practically as far as the layer where the surface with vanishing temperature and density would be found. Numerical calculations are made for the brighter component of Capella.

The author is very much indebted to Professor Harlow Shapley for giving him the opportunity of working at the Harvard Observatory.

CRYSTALLINE PROTEIN MOLECULES*

BY

Edwin J. Cohn, I. Fankuchen, J. L. Oncley, H. B. Vickery,
AND B. E. Warren

CONTENTS

INTRODUCTION TO THE CONFERENCE ON CRYSTALLINE PROTEIN MOLECULES

BY

By Edwin J. Cohn

*From the Department of Physical Chemistry, Harvard Medical School,
Boston, Massachusetts*

These conferences had their origin, I am told, in the notion that despite the large number of scientific meetings and symposia there remained a need for informal critical discussion of the problems confronting the active investigators in any given field. During the first year of this section, 1938–1939, I attended two of the three conferences. The first was on Electrophoresis and had been arranged by Dr. Duncan A. MacInnes, the third was on Dielectrics and had been arranged by Dr. Charles P. Smyth. In both there was an impressive balance between experiment and theory. In both there was incisive comment as a result of which more than one new investigation has been undertaken in order to test tentative hypotheses put forward as possible explanations of the accumulated evidence.

In both the conference on Electrophoresis and in that on Dielectrics the unifying principles were theoretical and technical. The laws of electrophoresis were considered in terms of all manner of ions, or of particles that are electrically charged. The application of dielectric theory was considered both to insulating materials and to biochemical systems. For the very nature of fundamental science is that laws developed for one kind of material or system have validity that transcends the origins of these laws.

In both the conference on Electrophoresis and the conference on Dielectrics reference was repeatedly made and deductions were drawn from investigations upon substances which, though diverse, in nature and function, form a well defined class of molecules; the proteins. The importance of these substances for biology and medicine; for agriculture and industry; is such that they have in recent years been investigated by many and ever increasingly powerful tools. In the course of these conferences it occurred to us, however, that the very specialized nature of the tools might well result in the future in the training of men who knew the tools more intimately than the substances they were investigating. Indeed, each tool begins to have so interesting a history and to involve so intricate a theory that few not

trained in its use can interpret the important investigations that are being reported and critically appraise the evidence that may be deduced from them.

The approaches to protein chemistry have in recent years been with the tools of the physicist and the physical chemist as well as with those of the organic chemist who yielded us our first notion of the nitrogenous nature of the proteins, of their composition in terms of amino acids and of the manner in which the amino acids were bound together in polypeptide linkage. The development through organic chemistry, begun in the nineteenth century, has continued in the twentieth century with the discovery of new amino acids and improvement in the methods of hydrolysis and of the isolation of amino acids from the protein hydrolysates. These developments have led to increases in our knowledge of the amino acid composition of proteins and therefore, in the ratio of the chemical groups of different kinds that are free in each protein.

There would appear to be a growing conviction that differences in the properties of proteins, as a result of which some are elements of structure in biochemical systems, others enzymes, hormones or viruses, may inhere in the distribution and spatial arrangement of these chemical groups of diverse nature. The notion of distribution involves notions of size and shape and the most potent tools yet available for the investigations of these properties of proteins have been developed by physicists and physical chemists.

The ultracentrifuge, developed by Svedberg, is by far the most powerful tool that we have for the study of the mass of the protein molecule. It has confirmed the idea that proteins are among the largest molecules known and has largely supplanted the earlier methods of estimating molecular weight by osmotic pressure or by ultrafiltration as a multiple of an analytically deduced minimal molecular weight. The calculation of molecular weight from the sedimentation velocity of a molecule in the ultracentrifuge generally involves knowledge of its free diffusion. Diffusion measurements, if carried out with sufficient accuracy, would of themselves yield molecular weights if the molecules were spherical. Since most protein molecules appear not to be spherical, a combination of diffusion and sedimentation velocities has been employed in calculating molecular weights; of molecular weights and diffusion constants in calculating asymmetry or of diffusion and viscosity—a property independent of size for incompressible, uncharged, spherical molecules—in the calculation of molecular weight.

Protein molecules may be oriented if they are asymmetrical either with respect to shape, or to the distribution of their electrically charged groups. In the former case they reveal double refraction of flow, in the latter they increase the dielectric constant of the solvent. The amount of the latter effect, yields, as a vector sum, the dipole distance and moment of all the electrically charged groups of the protein. The frequencies at which dispersion of the dielectric constant occurs, and the shape of the curve, can be employed in the calculation of protein asymmetry if the molecular weight is known since the time of relaxation of an oriented molecule is greater the greater its asymmetry. The relaxation times and the dimensions of proteins can be estimated both from measurements of the dielectric constant and of double refraction of flow, and compared with those derived from diffusion and viscosity.

Each of these measurements, when carried out with adequate technical skill, yields data of the utmost importance in the characterization of proteins; provided these have been so purified that only one kind of molecular species is investigated. Some of these techniques themselves yield evidence as to whether or not a given protein preparation consists of more than one molecular species. The interpretation of results when protein mixtures are investigated is far more difficult, however, than when only a single kind of chemical individual is studied. For the purposes of the present conference we propose therefore to limit discussion to proteins which have thus far been isolated as relatively pure chemical individuals. The criteria of chemical individuality are thus also worthy of discussion. Thus, crystallization is unquestionably a great aid in purification, but crystallization, even five recrystallizations (as in the case of the serum albumins of the horse), is not adequate proof that a chemical individual has been isolated, nor is electrophoretic mobility, dielectric dispersion, nor sedimentation velocity.

Protein molecules of the same family are often closely related and solubility measurements often reveal the existence of mixtures in systems even when all of the molecules have the same size, shape, and net charge. A protein whose solubility is independent of the amount of the solid crystalline phase with which it is in equilibrium may be considered a chemical individual. The composition of every preparation of such a protein should be the same. The obstacles to deducing composition, let alone structure, from analytical measurements upon the amino acids of a mixture of proteins are, however, almost insuperable.

And yet sufficiently precise analytical measurements upon those

amino acids present in a pure protein in small amount should be of the greatest value in determining the minimal molecular weight and thus aiding in the calculation of its true molecular weight. There should thus be a symbiotic relation between the organic chemist and the physical chemist in their attack upon the structure of proteins. In the isolation of a protein as a chemical individual the tools both of analytical and physical chemistry are of inestimable value. Once the protein is isolated and of proven purity, the organic chemist should be able to supply information regarding the nature of the groups that are present and something regarding their relations to each other which should be as helpful to the physical chemist as are his measurements to the analytical chemist.

Three examples of interrelations between the analytical and physical chemistry of the proteins may be cited. The determination of the iron in hemoglobin early led to an accurate estimation of the minimal molecular weight of this protein. Measurements of sulfur and sulfide sulfur led T. B. Osborne to consider the molecular weights of many proteins in terms of this element, largely present as the amino acids methionine and cystine or cysteine. The latter amino acid is present in very small amount in many proteins, as are tryptophane and tyrosine, and the analyses available at that time led us, some fifteen years ago, to attempt a systematic evaluation of the minimal molecular weights of the proteins in terms of their amino acid compositions. Newer and more accurate analytical methods are now available for these and other amino acids and a far more accurate estimate of the minimal molecular weights of those proteins, which have been proven to be chemical individuals, should thus now be possible. I should rather see the analytical results employed in the calculation of minimal molecular weights, which they are capable of yielding without further assumption, than interpreted in terms of molecular weights determined by physical-chemical considerations. The results from the two independent approaches may thus be compared, the errors in each procedure be more accurately appraised, and the cumulative evidence lead to an overwhelming confidence in our methods.

A second example of the comparison between the results of analytical and physical chemical studies of proteins could be derived from the estimates that have been made, on the one hand, of the acid-combining capacity of proteins, and on the other, of the dibasic amino acids, especially histidine, arginine and lysine. There would appear to be little doubt that, with the possible exception of a few terminal a-amino groups, it is the ϵ-amino, the imidazole and the guanidine groups of

proteins that combine with acids. In the sixteen or more years since we first became interested in this relation, though the results of the physical chemist have been amplified or changed in fewer instances than have those of the analytical chemist, the correspondence of the evidence from these totally different methods has come ever closer.

These basic groups as well as the carboxyl, the hydroxyl and sulfhydryl, the methianyl, indole and benzene groups of proteins appear to be largely responsible for their chemical reactivity. The basic and carboxyl groups are generally charged in neutral solution. In so far as the charged groups lead to the amphoteric properties of the proteins, problems arise which are of the greatest interest, but which Dr. MacInnes and I decided to omit from this discussion in favor of a future conference on the net charge of the protein molecule.

In so far as these same groups give evidence regarding the composition and structure of proteins, however, they must be considered here. Even in the present incomplete state of knowledge regarding the amino acid composition of proteins several relations can be shown, among them that between the nature of the individual groups of the protein and its density. The results of many investigators, especially those of Svedberg and his co-workers and of Adair, demonstrate that the densities of all proteins in solution, referred to their dry weight, are close to 1.33; that is, their specific volumes are close to 0.75. Some years ago, Edsall, McMeekin, and I had occasion to investigate the specific volumes of a large number of amino acids and peptides, of their derivatives and isomers. Certain additive volumes were deduced for the groups that they contained and found to be consistent not only with each other, but with the results deduced by Traube, from the study of other organic molecules. Expressed as the specific volumes of certain of the groups of which proteins are composed, our results yield 0.420 for $COOH$, 0.465 for $CONH$, 0.481 for NH_2, and 1.163 for CH_2.

Despite the wide variation between these results, the specific volumes of amino acid residues vary from 0.59 for aspartic acid, 0.65 for glycine and glutamic acid, and 0.68 for histidine to 0.80 for lysine, 0.85 for valine and 0.89 for leucine. The specific volumes of all proteins must fall within these limits. Those of proteins which yield large amounts of glycine and aspartic or glutamic acid on hydrolysis should be smallest, those containing large amounts of leucine or lysine largest. Even with the incomplete yields of amino acids that have thus far been isolated by the organic chemist, a specific volume of 0.74 may be calculated for egg albumin, as compared with the observed value of 0.749, for insulin of 0.74 as compared with 0.749 observed and for edestin of 0.73 as compared with 0.744 observed.

Increase in our knowledge of the composition of proteins on the one hand, and of their partial specific and molal volumes on the other, should bring even closer the results from this and kindred relations between organic and physical-chemical evidence.

During the period of this slow but steady progress in relating the properties of proteins to their composition there have been a succession of interesting theories of protein structure. The one in fashion when I was a student had the hexone bases at the center of the molecule, the mono-amino mono-carboxylic acids radiating from them. I still remember a beautiful stellate model of a protein that hung in T. B. Osborne's laboratory when I was a student there. It stimulated talk and suggested new experiments, especially on the hexone bases, and resulted in improvements in the analytical procedures for these amino acids.

Then came a period of disruptive protein structure. The amino acids were said to be present as diketopiperazines and the molecular weights of the proteins—in phenol—to be about 300. But these low estimates were due rather to the water from which it is always difficult to free proteins, than to the proteins themselves.

We hold no final view regarding the structure of proteins. We have no notion as to whether they are best regarded as peptide chains wound in some regular pattern or as hollow cages. The evidence from physical chemistry connects their size and shape, and something regarding the number and distribution of their charged and reactive groups. We know that few proteins are spherical, that they are hydrated and compressible; and that most are readily denatured, either with increased, decreased, or unchanged molecular weight.

The ultracentrifuge was early recognized as the most potent tool with which to study the sizes of proteins, and at an even earlier date the X-ray had been suggested as the most potent tool in the study of the structure of proteins. I remember arriving in Cambridge, England, in 1920 to study the work and methods of William Hardy, Frederick Hopkins, and Joseph Barcroft, and being shown by the latter a list of projects that had been drawn up, as I remember it, by the Medical Research Council, or some comparable body. On it was a suggested study of the X-ray diffraction of hemoglobin. That was twenty years ago this spring. There has since been an X-ray study of hemoglobin, and this study is, I understand, still continuing under Bragg's direction at the Cavendish Laboratory in Cambridge.

The first X-ray picture of a protein I ever saw was of edestin and had been made at the Kaiser Wilhelm Institute in Dahlem in 1926.

At this time a great deal of work was being carried out on the X-rays of fibers, especially cellulose fibers. Meyer and Mark, who were concerned with these studies in the laboratories of the I. G. Farbenindustrie also examined some protein fibers and had a study made of the dimensions of the glycine molecule by X-ray methods by Hengstenberg and Lenel in their laboratories. As far as I know, this was the first study of an amino acid but the need to determine the dimensions of the amino acids of which proteins are composed was recognized by Bernal, who early carried out a survey, and is now being reconsidered by Pauling, Corey, and their co-workers.

It is not my purpose in these introductory remarks to impinge upon the contributions of those who will come after me and who will have far more competence than I to discuss the X-ray work on proteins. It is but natural that so much of this work has been done in England, and Astbury's work, especially on the fibrous proteins, and that of Bernal, Miss Crowfoot, Fankuchen, and others on protein crystals, is beginning to yield a great deal of information regarding a number of proteins. In how far the evidence derived from such studies can be employed in estimating the number of molecules per unit cell, the dimensions of the molecules and important elements in their structure will, I hope, emerge from this discussion.

If the amino acid analysis of the organic chemist yields the minimal molecular weight, the X-ray analysis of the physicist yields the weight—a maximum weight—of the number of molecules in a unit cell. The molecular weight in solution, though it will sometimes equal either the minimum or maximum value, will generally be an intermediate integer. In terms of the molecular weights alone there is thus the possibility of relating the physical, physical-chemical, and organic-chemical evidence.

The X-ray photographs of the same protein when it is wet and after it has been dried reveal certain changes probably produced by the removal of water from the crystals. Is the water estimated in this way related to the water that Sorensen long since determined to be present in crystalline egg albumin and to the water of hydration that Adair and his co-workers have estimated to be present from a variety of methods? What volume of water is subjected to electrostriction by the charged groups of the protein? In how far will X-ray studies tell us how this water is distributed and how it is held?

The interpretation of X-ray patterns has become a vastly complex, technical subject and we shall hope to hear in this discussion how far the structure can be deduced from X-ray studies alone, how far this

information is independent of, how far supplemental to knowledge of proteins gained by other physical, physical-chemical, and organic-chemical methods. In the end, the X-ray pattern must depend upon the arrangements of the same groups that the organic chemist tells us are present and any distribution of these groups suggested by physical-chemical measurements must also be consistent with deductions from X-ray analysis. Indeed, it would appear an inevitable conclusion that sooner or later the evidence from these three different approaches to protein chemistry must yield a consistent and far more complete picture of the protein molecule than is now available. It was our thought in arranging for this conference that the time might well be at hand to explore the present state of our knowledge; to discover the nature of the assumptions and the nature of the evidence on the basis of which new hypotheses may be formulated and new investigations planned.

EVIDENCE FROM ORGANIC CHEMISTRY
REGARDING THE COMPOSITION
OF PROTEIN MOLECULES

By H. B. Vickery

From the Connecticut Agricultural Experiment Station, New Haven, Connecticut

The study of proteins has undergone a very interesting sequence of changes when considered from the standpoint of the scientific background of the workers in the field. The earliest contributors were for the most part affiliated with medicine. Beccari (1682–1766), who in 1747 prepared the first protein of vegetable origin, was both physician and natural philosopher. At various stages in his career he taught chemistry, physics, mathematics, logic, anatomy, and medicine, and made definite contributions in the fields of human nutrition, of weather observation and of disease. Fourcroy (1755–1809), to whom we owe the terms albumin and gelatin, was a distinguished physician and organizer of medical education. Scheele (1742–1786), who carried out the first noteworthy chemical investigation of casein, was an apothecary. Braconnot (1780–1854), who discovered glycine and leucine in 1820, was also at one stage in his career a pharmacist. Wollaston (1766–1828), the discoverer of cystine, was a physician and physicist as well as chemist.

From such investigators one would expect shrewd factual observation but not analysis nor theory and, even in the next stage, when organic chemists assumed the burden of investigation, attempts to account for observations were rare. Berzelius supplied nomenclature (cystine, glycine); Mulder coined the word protein and advanced the first speculation on what these substances were, thereby stimulating much research; Liebig discovered tyrosine and taught Ritthausen, the first of the great protein chemists.

The period from 1820 to 1900 was one in which organic chemists and those whose interest in physiology converted them into the first generation of biochemists, such men as Kühne, Hofmeister, Kossel, and Schulze, advanced our knowledge enormously. They laid a foundation that made the emergence of a Fischer inevitable, and perhaps their greatest contribution was the discovery of the various amino acids of which the protein molecule is formed. Under the spur of Fischer's formulation of an acceptable protein theory, as well as of his own discovery of no less than three protein amino acids, progress

became increasingly rapid. New types of investigation were undertaken which have led logically to the development of the present phase of protein chemistry in which physical chemists are playing a most important role.

It is my assignment, in the present conference on general protein chemistry, to outline some of the major contributions of the organic chemist. It seems to me that the most important of these contributions has been the concept that proteins are, from the point of view of structure, composed of amino acids. With respect to the details of this structure we *know* relatively little, although this is a theme that occupies many minds at the present time. Regardless of details of how they are combined with each other, however, it is fundamental to know, first, what are these amino acids, and, second, what are the proportions in which they are found in the protein molecule. It is with these two questions that I propose to occupy your time.

AMINO ACID THEORY OF PROTEIN CONSTITUTION

We are today so accustomed to the view that amino acids, in some form of combination, make up the bulk, at any rate, of protein molecules that few stop to inquire how this notion first became current. It is, of course, a notion implicit in the earliest work, that of Braconnot and of Liebig and his student Bopp, but I think it received its earliest explicit statement in Ritthausen's great book on "The Proteins of the Cereals, Legumes and Oil-Seeds," published at Bonn in 1872. The third section of this book is entitled "The decomposition products of gluten proteins, of legumins, and of conglutins", these being three classes of proteins he had recognized and differentiated, and in the first paragraph he points out that glutamic and aspartic acids are yielded in different proportions by different proteins and accordingly may serve "for the characterization of protein substances and as a foundation for the recognition of their individuality." This is a precise statement of the purpose of amino acid analysis and a recognition of its fundamental significance. Later in this chapter (on p. 222), he gives the first table of the amino acid composition of proteins to be published. It is very simple—only two amino acids are included—but the germ of the idea is there.

Within a year, and so far as I can tell quite independently, Hlasiwetz and Habermann[1] arrived at a similar conclusion and published the first amino acid analysis of casein. They did not give quantitative values

---

[1] **Hlasiwetz, H., & Habermann, J.** Jour. prakt. Chem. **7**: 397. 1873; Ann. **169**: 150. 1873.

but they maintained that glutamic acid, aspartic acid, leucine, tyrosine, and ammonia account for practically the whole of the casein molecule.

This view was not shared by all. Schulze grasped it, as did Kossel, Fischer, and Hofmeister, but, to the great Kühne and his school, amino acids were quite unimportant and Kühne's pupil Neumeister, in the 1897 edition of his textbook, devotes amazingly little attention to them. Tyrosine, leucine, and aspartic acid are regular digestion products; glutamic acid can be formed by the action of strong acids; glycine, alanine, and serine have some significance in connection with silk and that is all. It required Hofmeister, Fischer, and Kossel to make the significance of amino acids clear.

Today we are faced with a new problem, or rather with the modern development of this same old problem. In its present-day dress, this problem takes the form of the question, *which* are the amino acids that are of significance in protein chemistry? There is universal agreement that certain amino acids are of importance, and there is universal agreement that certain others are not. But there is a most interesting group of these substances about which opinions may and do differ widely, and it seems to me that we may well pause to inquire what is required by way of evidence before a given substance may be, as it were, admitted into this inner circle of the elect.

CRITERIA FOR ACCEPTANCE OF AN AMINO ACID

Some years ago Schmidt and the author[2] somewhat arbitrarily set up certain criteria by which the validity of a claim for the presence of any given amino acid among the products of hydrolysis of proteins might be tested. Our motives were thoroughly practical and entirely selfish. We were considering each authentic amino acid in detail and we did not wish to encumber our discussion with a good deal of material that we were convinced was rubbish. On the other hand, there were several cases—borderline cases if you will—where the evidence appeared almost adequate but that we felt required further study. In order to avoid saying with such grace as we could that we did not accept the data as published, we decided that, in order to be accepted, an amino acid must also have been isolated by some worker other than the discoverer and that its constitution must have been established by suitable synthetic means. Obviously, too, it must have been derived from a pure protein by hydrolysis, and must be thoroughly characterized by the preparation of salts and other derivatives. It was desirable, as further conclusive evidence, that its utilization by the animal body should have been demonstrated.

[2] Vickery, H. B., & Schmidt, C. L. A. Chem. Rev. 9: 169. 1931.

TABLE 1

CLASSIFICATION OF AMINO ACIDS

A. Amino acids concerning which there is no doubt whatever, arranged in chrono-
logical order of discovery as products of protein hydrolysis.

1. Glycine[1]	10. Arginine[9]
2. Leucine[1]	11. Histidine[10, 11]
3. Tyrosine[2]	12. Valine[12]
4. Serine[3]	13. Proline[12]
5. Glutamic acid[4]	14. Tryptophane[13]
6. Aspartic acid[5]	15. Hydroxyproline[14]
7. Phenylalanine[6]	16. Isoleucine[15]
8. Alanine[7]	17. Methionine[16]
9. Lysine[8]	18. Threonine[17]

B. Amino acids that occupy a special position because of their narrow range of
distribution or for other reasons.

19. Thyroxine[18] (thyroid gland proteins)
20. Diiodotyrosine or iodogorgoic acid[19] (thyroid gland protein and skeleton
protein of certain marine organisms)
21. Dibromotyrosine[20] (skeleton of *Primnoa lepadifera*)
22. Norleucine[21] (spinal cord protein)
23. Cystine[22] (universally distributed; mode of linkage of sulfur still in debate)
24. Cysteine[23] (evidence from nitroprusside reaction of few native proteins)
25. Hydroxyglutamic acid[24] (existence in proteins doubted by some investi-
gators)

C. Amino acids known as plant constituents that may possibly be expected to be
found in proteins.

1. Thiolhistidine[25] (in ergot as betaine ergothioneine; in blood)
2. Dihydroxyphenylalanine[26, 27] (in bean seedlings; probably widely dis-
tributed)
3. Citrulline[28] (in watermelon tissue; probably of metabolic significance in
urea formation in animals)
4. Canavanine[29] (in certain beans)
5. Djenkolic acid[30] (in Djenkol bean)

D. Amino acids for which claims have not been substantiated.

6. Amino butyric acid[31]
7. Hydroxyvaline[32]
8. Hydroxylysine[33]
9. Norvaline[34, 35]
10. Diaminoglutaric acid[36]
11. Diaminoadipic acid[36]
12. Hydroxyaspartic acid[36]
13. Dihydroxydiaminosuberic acid[36]
14. "Caseianic acid"[36]
15. "Caseinic acid"[36]
16. Prolysine (α-amino-ϵ-hydantoin caproic acid)[37]
17. Hyphasamine ($C_{16}H_{24}O_5N_2$)[38]

<div align="center">

TABLE 1—*Continued*

CLASSIFICATION OF AMINO ACIDS

</div>

18. Dodecandiaminodicarboxylic acid[39]
19. Base $C_4H_{11}O_3N$ [10]
20. Protoctine $(C_8H_{15}O_4N_2)$ [41]
21. Diaminotrihydroxydodecanic acid[42, 43]
22. Hydroxytryptophane[44, 45]

[1] Braconnot, H.   Ann. chim. phys. (2) 13: 113.   1820.
[2] Bopp, F.   Ann. 69: 16.   1849.
[3] Cramer, E.   Jour. prakt. Chem. 96: 76.   1865.
[4] Ritthausen, H.   Jour. prakt. Chem. 99: 454.   1866.
[5] Ritthausen, H.   Jour. prakt. Chem. 103: 233.   1868.
[6] Schulze, E., & Barbieri, J.   Ber. 14: 1785.   1881.
[7] Weyl, T.   Ber. 21: 1407.   1888.
[8] Drechsel, E.   Jour. prakt. Chem. 39: 425.   1889.
[9] Hedin, S. G.   Zeit. physiol. Chem. 20: 186.   1895.
[10] Hedin, S. G.   Zeit. physiol. Chem. 22: 191.   1896-7.
[11] Kossel, A.   Zeit. physiol. Chem. 22: 176.   1896-7.
[12] Fischer, E.   Zeit. physiol. Chem. 33: 151.   1901.
[13] Hopkins, F. G., & Cole, S. W.   Proc. Roy. Soc. London 68: 21.   1901.
[14] Fischer, E.   Ber. 35: 2660.   1902.
[15] Ehrlich, F.   Ber. 37: 1809.   1904.
[16] Mueller, J. H.   Proc. Soc. Exp. Biol. & Med. 19: 161.   1922.
[17] Meyer, C. E., & Rose, W. C.   Jour. Biol. Chem. 115: 721.   1936.
[18] Kendall, E. C.,   Trans. Assoc. Am. Physicians, 30: 420.   1915.
[19] Drechsel, E.   Zeit. Biol. 33: 96.   1896.
[20] Mörner, C. T.   Zeit physiol. Chem. 88: 138.   1913.
[21] Abderhalden, E., & Weil, A.   Zeit. physiol. Chem. 81: 207.   1912.   84: 39.   1913.
[22] Mörner, K. A. H.   Zeit. physiol. Chem. 28: 595.   1899.
[23] Mirsky, A. E., & Anson, M. L.   Jour. Gen. Physiol. 18: 307.   1935.
[24] Dakin, H. D.   Biochem. Jour. 12: 290.   1918.
[25] Eagles, B. A., & Johnson, T. B.   Jour. Am. Chem. Soc. 49: 575.   1927.
[26] Torquati, T.   Arch. farm. sper. 15: 213.   1913.
[27] Guggenheim, M.   Zeit. physiol. Chem. 88: 276.   1913.
[28] Wada, M.   Biochem. Zeit. 224: 420.   1930.
[29] Kitagawa, M., & Tomita, T.   Proc. Imp. Acad. Japan 5: 380.   1929.
[30] van Veen, A. G., & Hyman, A. J.   Rec. trav. chim. 54: 493.   1935.
[31] Abderhalden, E., & Weil, A.   Zeit. physiol. Chem. 81: 207.   1912.
[32] Schryver, S. B., & Buston, H. W.   Proc. Roy. Soc. London B99: 476.   1925.
[33] Schryver, S. B., Buston, H. W., & Mukherjee, D. H.   Proc. Roy. Soc. London B98: 58.   1925.
[34] Abderhalden, E., & Bahn, A.   Ber. 63: 914.   1930.
[35] Abderhalden, E., & Reich, F.   Zeit. physiol. Chem. 193: 198.   1930.
[36] Skraup, Z. H.   Zeit. physiol. Chem. 42: 274.   1904.
[37] Wada, M.   Proc. Imp. Acad. Japan 9: 43.   1933.
[38] Engeland, R.   Biochem. Jour. 19: 850.   1925.
[39] Fränkel, S., & Friedmann, M.   Biochem. Zeit. 182: 434.   1927.
[40] Gortner, R. A., & Hoffman, W. F.   Jour. Am. Chem. Soc. 47: 580.   1925.
[41] Schryver, S. B., & Buston, H. W.   Proc. Roy. Soc. London B100: 360.   1926.
[42] Fischer, E., & Abderhalden, E.   Zeit. physiol. Chem. 42: 540.   1904.
[43] Fischer, E.   Zeit. physiol. Chem. 99: 54.   1917.
[44] Abderhalden, E., & Kempe, M.   Zeit. physiol. Chem. 52: 207.   1907.
[45] Abderhalden, E., & Sickel, H.   Zeit. physiol. Chem. 144: 80.   1925.

By the use of these criteria, we eliminated from discussion a series of substances for which claims have been advanced from time to time. Some of these occupy an altered position today.

In TABLE 1, the amino acids are classified in four groups; in the first are eighteen substances about which there is no question whatever. These, at least, have been positively identified among the products of hydrolysis of proteins and, moreover, most of them may be expected to be present in larger or smaller amount in the hydrolysate of any given protein.

A second group has been formed of amino acids which are either rarely encountered because of a narrow range of known distribution, or because certain details of their mode of combination in the protein molecule are still in debate.

A third group includes amino acids known in nature as plant constituents and therefore of significance in amino acid metabolism. None of these substances has been as yet recognized as a product of protein hydrolysis.

In a fourth group are placed substances, or preparations, which figure in the literature as protein amino acids but concerning which more or less doubt is still entertained.

In the following paragraphs, a number of these amino acids are discussed in some detail in order to show what may be regarded as the present position of our knowledge of them.

Amino Butyric Acid

Claimed by Schutzenberger and Bourgeois[3] in 1875, by Foreman[4] in 1913, by Abderhalden and Weil[5] in 1912, and by Abderhalden and Bahn[6] in 1937, but insufficiently characterized.

Hydroxyamino Butyric Acid

Claimed by Schryver and Buston in 1925, by Gortner and Hoffman[7] in 1925, and Rimington[8] in 1927. None of these preparations was completely characterized, and the position of the hydroxyl group was not established. In 1935, Rose and his collaborators[9] demonstrated conclusively that this substance is present in several proteins, notably casein, and have since established its constitution by synthesis and

[3] Schutzenberger, P., & Bourgeois, A.   Compt. rend. 81: 1191.   1875.
[4] Foreman, F. W.   Biochem. Zeit. 56: 1.   1913.
[5] Abderhalden, E., & Weil, A.   Zeit. physiol. Chem. 81: 207.   1912.
[6] Abderhalden, E., & Bahn, A.   Zeit. physiol. Chem. 245: 246.   1937.
[7] Gortner, R. A., & Hoffman, W. F.   Jour. Am. Chem. Soc. 47: 580.   1925.
[8] Rimington, C.   Biochem. Jour. 21: 1187.   1927.
[9] Meyer, C. E., & Rose, W. C.   Jour. Biol. Chem. 115: 721.   1936.

named it *threonine*. It is now regarded as thoroughly established as a protein decomposition product.

Hydroxyvaline

Claimed by Schryver and Buston in 1925, by Czarnetzky and Schmidt[10] in 1931 and by Brazier[11] in 1930. Abderhalden and Heyns[12] discussed this substance in 1934 and expressed serious doubt of its existence in proteins.

Hydroxylysine

Claimed by Schryver and his associates in 1925 and 1927, and by Van Slyke and his associates[13] in 1938 as a constituent of gelatin, but not yet sufficiently characterized and its constitution established.

Norvaline

Claimed by Abderhalden and Bahn in 1930 and also by Abderhalden and Reich the same year. In 1932, Abderhalden and Heyns[14] advanced further evidence for their claim. Doubt of the existence of this substance among the products of hydrolysis of proteins may still quite properly be entertained since it has been secured in only one laboratory; nevertheless Abderhalden's evidence is quite extensive.

Norleucine

Claimed by Thudichum in 1901, and by Abderhalden and Weil in 1912[5, 15] as a constituent of spinal cord protein. More recently it has been claimed by Yaginuma, Arai, and Hayakawa[16] also in spinal cord protein, by Nuccorini[17] in castor bean protein, and by Abderhalden and Heyns.[18] Czarnetzky and Schmidt[19] have confirmed much of Abderhalden's early work and advance evidence very little if at all short of being completely conclusive that this substance is present among the products of hydrolysis of spinal cord protein. The difficulty is to distinguish between the physical properties of leucine, isoleucine, and norleucine. Small differences only are to be expected and separations are accordingly unusually difficult. However, many if not most students are prepared to accept the evidence today.

[10] **Czarnetzky, E. J., & Schmidt, C. L. A.** Jour. Biol. Chem. 92: 453. 1931.
[11] **Brazier, M. A. B.** Biochem. Jour. 24: 1188. 1930.
[12] **Abderhalden, E., & Heyns, K.** Zeit. physiol. Chem. 229: 236. 1934.
[13] **Van Slyke, D. D., Hiller, A., Dillon, R. T., & MacFadyen, D.** Proc. Soc. Exp. Biol. Med. 38: 548. 1938.
[14] **Abderhalden, E., & Heyns, K.** Zeit. physiol. Chem. 206: 137. 1932. 209: 27. 1932.
[15] **Abderhalden, E., & Weil, A.** Zeit. physiol. Chem. 84: 39. 1913.
[16] **Yaginuma, T., Arai, G., & Hayakawa, K.** Proc. Imp. Acad. Tokyo 8: 91. 1932.
[17] **Nuccorini, R.** Ann. chim. appl. 24: 25. 1934.
[18] **Abderhalden, E., & Heyns, K.** Zeit. physiol. Chem. 214: 262. 1933.
[19] **Czarnetzky, E. J., & Schmidt, C. L. A.** Jour. Biol. Chem. 97: 333. 1932

3, 5-Dibromotyrosine

This was isolated by Mörner[20] from the skeleton of a coral unusually high in bromine. A survey of a number of marine organisms showed that bromine is widely distributed, although usually present in only small amounts, and the richest species was investigated. The validity of Mörner's observation has never been called in question although, so far as I know, it has not been repeated. Dibromotyrosine may probably be regarded as a well-established amino acid of very limited known distribution and thus occupies a position like that of norleucine. The iodine analogue is probably widely distributed in marine organisms and in mammalian thyroid tissue.

Dihydroxyphenylalanine

This substance is known in nature in the seed pods of certain legumes and is also recognized as a product of enzymatic oxidation of tyrosine.[21] It occupies an important position in theories of melanin formation in plant tissues and animals. So far it has not been observed among the products of hydrolysis of proteins but is by no means an improbable component.

Thiolhistidine

Represented in nature by the betaine ergothioneine found in ergot by Tanret, and is now known to be a constituent of the blood of animals fed a corn diet.[22] The amino acid has been sought for in zein, but, save that hydrolysates of this protein give what is held by Hunter to be a specific color test for the thiolimidazole ring, without success.

Canavanine

Isolated by Kittagawa and Tomita[23] in 1929 from extracts of the jack bean, it was found to be a basic substance of the formula $C_5H_{11}N_4O_3$ and was later shown[24] to be derived from γ-hydroxy-α-aminobutyric acid. Its constitution was established by Gulland and Morris[25] as a hydroxyguanidino derivative of this acid, thus confirming the structure assigned by the Japanese workers. Its significance in protein chemistry has yet to be established.

[20] Mörner, C. T. Zeit. physiol. Chem. 51: 33. 1907. 88: 138. 1913.
[21] Evans, W. C., & Raper, H. S. Biochem. Jour. 31: 2155. 1937.
[22] Eagles, B. A., & Vars, H. M. Jour. Biol. Chem. 80: 615. 1928.
[23] Kitagawa, M., & Tomita, T. Proc. Imp. Acad. Tokyo 5: 380. 1929.
[24] Kitagawa, M., & Mononobe, S. Jour. Agr. Chem. Soc. Japan 9: 845. 1933.
[25] Gulland, J. M., & Morris, C. J. O. R. Jour. Chem. Soc 763. 1935.

Citrulline

In 1914, Koga and Odake[26] described the isolation of a substance $C_6H_{13}N_3O_3$ from the juice of watermelon. Aside from the fact that it formed a copper salt, little else was recorded. In 1930, Wada prepared the substance again and showed that its properties were best explained on the assumption that it is δ-carbamido ornithine. This was confirmed by synthesis.[27] Shortly afterwards Ackermann[28] repeated Wada's work and obtained citrulline not only from watermelon extract, but also from the products of the action of putrefactive organisms on arginine. Horn[29] likewise observed this last transformation although the yield was low.

Wada has also claimed[30] the isolation of citrulline from a tryptic digest of casein but the possibility of enzymatic production from arginine was not excluded.

Fearon[31] has recently described a color test which permits discrimination between citrulline and all of the well-known protein amino acids. This test is given by many proteins (egg albumin, casein, fibrin, etc.) and is regarded as evidence that citrulline exists preformed in the protein molecule. Incontrovertible evidence from isolation has not yet, however, been secured.

Djenkolic Acid

Prepared in 1935 by van Veen and Hyman[32] from extracts of the tropical Djenkol bean, it was held to be the cysteine thioacetal of formaldehyde, a substance which occupies an interesting relationship to both cystine and methionine. This structure was confirmed by du Vigneaud and Patterson[33] by synthesis and demonstration of identity with van Veen's own preparation. The possibility that djenkolic acid may be found among the products of hydrolysis of proteins has not as yet been excluded, but the evidence available suggests that it can be expected only in rare cases.

Ornithine

Reference should also be made to ornithine, the next lower homologue of lysine. The δ-guanidino derivative of this substance is the

[26] Koga, Y., & Odake, S. Jour. Tokyo Chem. Soc. 35: 519. 1914.
[27] Wada, M. Proc. Imp. Acad. Tokyo 6: 15. 1930. Biochem. Zeit. 224: 420. 1930
[28] Ackermann, D. Zeit. physiol. Chem. 203: 66. 1931.
[29] Horn, F. Zeit. physiol. Chem. 216: 244. 1933.
[30] Wada, M. Proc. Imp. Acad. Tokyo 8: 367. 1932.
[31] Fearon, W. R. Biochem. Jour. 33: 902. 1939.
[32] van Veen, A. G., & Hyman, A. J. Rec. trav. chim. 54: 493. 1935
[33] du Vigneaud, V., & Patterson, W. I. Jour. Biol. Chem. 114: 533. 1936.

well-known basic amino acid arginine, but so far as I am aware, ornithine has never been isolated from a protein except after treatment with arginase or with strong alkali. However, its isolation presents very special difficulties so that this is not in itself surprising.

Although not as yet to be classified as a protein amino acid, ornithine occupies an important position in certain theories of amino acid metabolism.[34]   The existence of an enzyme that decomposes arginine into urea and ornithine[35] is most significant for animal biochemistry, and the isolation of δ-N-acetyl ornithine from the roots of *Corydalis ochotensis*, a Siberian species of fumitory, by Manske[36] suggests that ornithine may play a role also in certain plants. The occurrence of ornithuric acid (dibenzoyl ornithine) in the excreta of birds to which benzoic acid has been administered has, of course, long been known.

Selenium Compound

It is perhaps of interest to mention a compound of selenium recently isolated by Horn and Jones[37] from plants grown in a selenium-containing soil. No information is yet available with regard to the source or the method of preparation, but the compound is held to be an amino acid and the structure provisionally assigned contains a cysteine and a homocysteine radical both attached to a single selenium atom. Whatever its structure, it is a natural amino acid of a new kind that will undoubtedly prove to be of unusual importance.

Other Compounds

In addition to these substances there is a long list of claims, some of which have been shown to be ill-founded, others of which still stand in the literature. Such substances as *diaminoglutaric* acid, *diaminoadipic* acid, *hydroxyaminosuccinic* (hydroxyaspartic) acid, *dihydroxydiaminosuberic* acid, claimed by Skraup in 1904, together with two additional vague substances called *"caseianic"* acid and *"caseinic"* acid, probably represent little more than errors of judgment. Fischer contributed *diaminotrihydroxydodecanic* acid in 1904, but withdrew it in 1917. *Hydroxytryptophane* was claimed by Abderhalden and Kempe in 1907, but was later found to have been a mixture of a peptide of tyrosine and proline with a little tryptophane.

Still other unconfirmed reports are those of Wada[38] who claims to

[34] Krebs, H. A., & Henseleit, K.   Zeit. physiol. Chem. **210**: 33.   1932.
[35] Kossel, A., & Dakin, H. D.   Zeit. physiol. Chem. **41**: 321. 1904.   **42**: 181.   1904.
[36] Manske, R. H. T.   Canadian Jour. Res. **B15**: 84.   1937.
[37] Horn, M. J., & Jones, D. B.   Jour. Am. Chem. Soc. **62**: 234.   1940.
[38] Wada, M.   Proc. Imp. Acad. Tokyo **9**: 43.   1933.

have prepared *α-amino-ε-hydantoin capronic* acid which he named prolysine, from gelatine. Wada[30] has also claimed the isolation of *citrulline* from proteins. Engeland[39] in 1925 obtained a material from elastin of the composition $C_{16}H_{24}O_5N_2$ that he called *hyphasamine* and regarded as a peptide of tyrosine with a hitherto unknown amino acid.

Fränkel and Friedmann[40] in 1927 believed they had obtained a *dodecandiaminodicarboxylic* acid from casein, and two years later Fränkel and Monasterio[41] claimed still another new substance that they had obtained from hemoglobin. Gortner and Hoffman[7] in 1925 prepared from teozein a basic substance to which the formula $C_4H_{11}O_3N$ was ascribed. Schryver and Buston[42] the following year obtained a substance $C_8H_{15}O_3N_3$ from castor bean protein that they named *"protoctine"*.

Even this list is probably incomplete, but sufficient has been included to illustrate the point I wish to make, namely that the isolation of a new amino acid is a matter that calls for a degree of self-criticism, in addition to skill, patience, and chemical insight, that is extremely rare. I sincerely believe that announcements of new amino acids must be treated with the utmost conservatism. Doubtless it is difficult if not impossible to set up rigid criteria for acceptance, but the very fact that about as many substances have been claimed as new protein amino acids in the past thirty-five years as have been satisfactorily established in the past one hundred and twenty shows the danger to which students are exposed. I know of very few textbooks that contain a list of protein amino acids that is above criticism. Even Schmidt's recent handbook omits threonine from the list of accepted amino acids and includes it in the group of those reported but not verified, although norleucine appears in the accepted list. To my mind the evidence for threonine so far outweighs that for norleucine as to leave no room for argument.

Two serious problems remain to be discussed in this connection: these are the position of cystine and cysteine, and the position of hydroxyglutamic acid.

Cystine is a disulfide and, according to the views of protein constitution advanced by Astbury, and by Mirsky and Pauling, represents a configuration within the molecule whereby two peptide chains, or possibly two folds of the same chain, are linked together through the

[39] Engeland, R.   Biochem. Jour. 19: 350.   1925.
[40] Fränkel, S., & Friedmann, M.   Biochem. Zeit. 182: 434.   1927.
[41] Fränkel, S., & Monasterio, G.   Biochem. Zeit. 213: 65.   1929.
[42] Schryver, S. B., & Buston, H. W.   Proc. Roy. Soc. London B100: 360   1926.

sulfur.   I do not know just how cystine is dealt with in the formula-
tions of Wrinch, but assume that it may also serve as some form of
linkage between adjacent fabrics.   On any view, what we isolate as
cystine is really present in the protein molecule fundamentally as
cystein and, regardless of any assumptions of protein structure, there
is considerable evidence today that cysteine does occur in certain
proteins. A few proteins give a nitroprusside reaction (myosin)
while in the native condition; many if not most proteins give this
test after denaturation.[43] We have no theory to account for the
presence of sulfhydryl groups in native proteins unless it be assumed
that cysteine forms one of the amino acids in the structure.   The pres-
ence of increased amounts of sulfhydryl after denaturation, or the
appearance of these groups through the agency of denaturation[44] would
appear to suggest that a rearrangement of the cysteine side chains oc-
curs whereby free sulfhydryl groups are liberated.   It is usually as-
sumed that much of the sulfhydryl is in combination in pairs in the
form of cystine, these linkages being broken by the rearrangement
that occurs when the protein is denatured.   Thus it seems probable
that both cysteine and cystine should be regarded as being present in
the original protein molecule, at least in many cases.   Certainly the
study of the relationships of these two substances in the molecule is
one of the most promising lines of attack upon the fundamental
problem of structure.

The other problem—the position of hydroxyglutamic acid—presents
several puzzling features.   In 1918, Dakin[45] encountered this substance
as a product of hydrolysis of casein.   He described a method of isola-
tion, difficult to be sure, by which he obtained on one occasion a yield
of approximately 10 per cent; usually however only from 2 to 3 per
cent could be isolated.   The substance was characterized by analysis,
by means of the salts of metals, by its conversion into the correspond-
ing pyrrolidone carboxylic acid and by the product of oxidation, malic
semialdehyde.   Somewhat later,[46] he described its preparation from
gliadin and glutenin and also gave a method for its synthesis.

There have been several reports in the literature of subsequent iso-
lations of this substance although none of these is entirely convincing.
Onuki[47] obtained "shapeless crystals" which gave the tests described
by Dakin and also a similar product of oxidation.   Calvery[48] gave an

[43] **Greenstein, J. P.**   Jour. Biol. Chem. **125**: 501.   1938.   **128**: 233.   1939.
[44] **Mirsky, A. E., & Pauling, L.**   Proc. Nat. Acad. **22**: 439.   1936.
[45] **Dakin, H. D.**   Biochem. Jour. **12**: 290.   1918.
[46] **Dakin, H. D.**   Biochem. Jour. **13**: 398.   1919.
[47] **Onuki, M.**   Jour. Chem. Soc. Japan **43**: 737.   1922.
[48] **Calvery, H. O.**   Jour. Biol. Chem. **94**: 613.   1931.

analytical figure for egg albumin based upon an isolation experiment performed on the hydrolysate from 250 gm. of protein. The identification depended upon the neutralization equivalent, the silver content of the silver salt, and the nitrogen content of the strychnine salt. His preparation was not crystalline but was obtained as a white powder by the use of alcohol. Gulland and Morris[49] obtained a small quantity from the barium sulfate precipitated during the removal of the acid from a hydrolysate of casein. The yield was very small, 0.33 per cent of the casein taken, but the substance was characterized as the nitrophenylhydrazone of malic semialdehyde.

Harington and Randall[50] have described a synthesis of β-hydroxyglutamic acid which makes the substance reasonably easily available. A careful comparison of the properties of the synthetic acid with those given by Dakin for the natural optically active substance showed many similarities and a few differences.

Synthetic (Harington)	Natural (Dakin)
Crystallizes from concentrated aqueous solution, usually in stout prisms with water of crystallization, $N = 8.5, 8.6\%$.	Crystallizes "slowly in thick prisms from its sirupy solution," $N = 8.40$, theory 8.59%. No mention of water of crystallization.
Oxidation with chloramine-T gave malic semialdehyde isolated as nitrophenylosazone and crystallized from nitrobenzene "in brownish red prismatic needles m.p. 291° (uncorr.)."	Oxidation with chloramine-T gave malic semialdehyde isolated as nitrophenylosazone and crystallized from nitrobenzene "in red-brown prismatic needles melting at 297–299°."
Silver salt insoluble, $N = 3.6$, theory 3.7%.	Silver salt insoluble, $N = 3.56$, Ag 57.3, theory 3.7 and 57.3%.

Certain differences were found by Harington and Randall; many of these, however, may well be due to stereoisomerism since four active and two inactive forms are theoretically possible. The synthetic acid crystallized from concentrated solution with three molecules of water of crystallization and was only moderately insoluble in cold water. When heated, the hydrated acid liquefied at 75°, solidified again and finally decomposed at 185°. The anhydrous acid decomposed at 198°. Dakin's product had no sharp melting point, and appeared to lose water slowly at 105° with ring closure. No such ring closure was

[49] Gulland, J. M. & Morris, C. J. O. R. Jour. Chem. Soc. 1644. 1934.
[50] Harington, C. R., & Randall, S. S. Biochem. Jour. 25: 1917. 1931.

observed with the synthetic acid even when heated at 110° over phosphorus pentoxide although this did occur when the aqueous solution was boiled at pH 4. Harington and Randall found that the hydrochloride of the synthetic acid has convenient properties for isolation and pointed out that its insolubility may account for some of the failures to isolate this substance that have been experienced. On the other hand, there is no evidence that the hydrochloride of the natural substance is notably insoluble.

Harington and Randall record complete failure of their attempts to isolate hydroxyglutamic acid from casein even when Dakin's directions were "meticulously followed." They state, however, that the fraction that should have contained this substance, although small and heavily contaminated with glutamic and aspartic acid, did in fact yield the oxidation product described by Dakin in "not inconsiderable amounts" and they did not hold the existence of hydroxyglutamic acid to be in doubt.

Nevertheless the repeated failures to secure evidence for the presence of hydroxyglutamic acid in protein hydrolysates that have been experienced have led to a feeling of uncertainty regarding this substance. Although, to my knowledge, no publications have appeared in which its occurrence is denied, it would seem desirable that this problem should be discussed.

I have one small contribution to make to this discussion. Several years ago Dr. James Melville, working in my laboratory, undertook to develop a new method to estimate hydroxyglutamic acid. It seemed possible that, if this substance were deaminized and then oxidized with permanganate and bromine under the same conditions as those employed for the oxidation of malic acid in the course of the analytical method for this substance,[51] a dinitrophenylosazone of the brominated oxidation product should be obtainable analogous to, if not identical with, the product from malic acid. This was found to be the case and a simple procedure to determine hydroxyglutamic acid (synthetic) was readily developed. Glutamic acid itself does not give this oxidation product but aspartic acid of course does. Accordingly it became necessary to develop a trustworthy procedure to separate glutamic and hydroxyglutamic acid from aspartic acid before the analytical method could be applied to protein analysis. The simplest way to do this seemed to be to convert the hydroxyglutamic and glutamic acids to their respective pyrrolidone carboxylic acids by ring closure; these

[51] Pucher, G. W., Vickery, H. B., & Wakeman, A. J.   Ind. Eng. Chem., Anal. Ed. 6: 288. 1934.

could then be extracted at a suitable reaction (about pH 3) by means of ethyl acetate in a continuous extraction apparatus of the Widmark type.[52] The extract would thus be freed from any contamination with aspartic acid, and the pyrrolidone carboxylic acids could then be hydrolyzed back to the straight chain compounds and the mixture analyzed by the oxidation procedure. Preliminary experiments only were completed before Dr. Melville had to return to New Zealand, but the observations he made were most encouraging from one point of view and most disturbing from another.

The conditions under which about 90 per cent of the acids could be converted to the respective ring compounds were found. It was sufficient to heat at boiling temperature in dilute solution for 24 hours. With protein hydrolysates, a shorter period of heating in the autoclave at 120–125° was adopted, but difficulty was experienced in obtaining so complete a ring formation; three successive heatings and extractions were necessary. The most serious trouble, however, was with the instability of the hydroxypyrrolidone carboxylic acid when attempts were made to hydrolyze it back to the straight chain compound. Significant losses were experienced even when 2N hydrochloric acid at boiling temperature for 1.5 hours was used; with 20 per cent acid the loss was serious.

This brought up the question whether hydroxyglutamic acid itself is stable when boiled with strong acid. When 1 gram of casein and 0.1 gram of hydroxyglutamic acid were boiled 26 hours with 30 per cent sulfuric acid, the pyrrolidone carboxylic acid recovered after the subsequent ring formation and extraction operations was only about 10 per cent of itself higher than that from a casein blank. The bromine-permanganate oxidation test was positive and its intensity corresponded to this small increase in pyrrolidone carboxylic acid as determined by the increase in amino nitrogen on hydrolysis. The inference was clear, however, that little hydroxyglutamic acid survived the treatment with boiling acid necessary for hydrolysis of the protein, and it was also significant that no oxidation test could be secured from hydrolysates of casein prepared in the usual way and subjected to the procedures for pyrrolidone carboxylic acid formation and extraction. This recalls the observation of Knoop and his associates[53]

[52] Pucher, G. W., & Vickery, H. B. Ind. Eng. Chem , Anal. Ed. **11**: 656. 1939. **12**: 27. 1940.
[53] Knoop, F., Ditt, F., Hecksteden, W., Maier, J., Merz, W., & Härle, B. Zeit. physiol. Chem. **239**: 30. 1936.

who found that synthetic hydroxyglutamic acid is in part decomposed to ammonia and α-ketoglutaric acid by boiling acid.*

These observations provide what is possibly a clue to the present situation with respect to hydroxyglutamic acid. Examination of Dakin's original paper shows that he hydrolyzed his casein for only 12 to 16 hours with 5 times its weight of 25 per cent by volume sulfuric acid, *i. e.* 35 per cent by weight. Is it possible that his success was due to a far less severe treatment of the protein than is now customary? Until these matters have been settled, I feel that one must go very carefully indeed before declining to accept Dakin's work.

One further point with regard to the amino acids that have been shown to be present in the protein molecule should perhaps be made. Aspartic and glutamic acids are indeed products of protein hydrolysis, and so fulfill the definitions that have been discussed above; but, in the protein molecule itself, these substances occur to a large extent as their respective amides asparagine and glutamine. The literature of amide nitrogen in proteins is long and interesting;[54] the suggestion that asparagine may be present in the intact protein molecule was made by Ritthausen in his book in 1872, and probably independently by Nasse the same year. The view was supported by the investigations of Hlasiwetz and Habermann which appeared the following year, and, in fact, the name glutamine was coined by these authors at that time, long before glutamine itself was recognized to be present in plant tissues by Schulze and ultimately isolated. In recent years, this early speculation has been confirmed by the isolation of asparagine from enzyme digests of edestin by Damodaran[55] and of glutamine from gliadin by Damodaran, Jaaback, and Chibnall.[56] Consideration of the quantitative relationships between the dicarboxylic amino acids and the amide nitrogen shows, however, that, in those proteins for which adequate data exist, only a part of the total amount of dicarboxylic acid can be present in amide combination. The exact distribution of the amide nitrogen between the aspartic and glutamic acid in any given case is a problem for the future.

DETERMINATION OF AMINO ACIDS

Sufficient has perhaps been said to indicate that no final answer can yet be given to the question, which are the protein amino acids? And,

---

* Dr. Dakin has informed me that he has been unable to convert natural hydroxyglutamic acid into α-ketoglutaric acid by this procedure.

[54] Vickery, H. B., & Osborne, T. B.   Physiol. Rev. 8: 393.   1928.

[55] Damodaran, M.   Biochem. Jour. 26: 235.   1932.

[56] Damodaran, M., Jaaback, G., & Chibnall, A. C.   Biochem. Jour. 26: 1704.   1932.

as I shall now show, a still less precise answer can be given to the question, how can the proportions of the different amino acids yielded by proteins best be determined? At the outset, the purpose of protein analysis must be defined. If one is merely interested in whether a series of protein fractions, secured by some systematic method of precipitation, differ from each other in chemical composition, there are literally scores of methods that may be employed. To show that the first and last members of the series differ in, for example, trypto-phane content, a notably inaccurate method is almost as good as a thoroughly accurate one, provided that it is applied under rigidly con-trolled conditions. However, if the purpose is to obtain the true tryptophane content so that calculation of molecular relations with other amino acids may be carried out, the choice of methods is neces-sarily more circumscribed. This is the type of analytical operation that is envisaged in the following discussion since I feel that we are now fundamentally interested in absolute values rather than in com-parative values.

For the purposes of discussion therefore, the well-known amino acids are divided into three groups, and I shall try to indicate the position of our analytical knowledge with respect to most of them.

It will be convenient to begin in reverse order and specify the amino acids for which there are no adequate analytical methods whatever. The substances in Group A of TABLE 2 are the amino acids for which our information is largely, if not entirely, qualitative. To be sure more or less of each of them has been isolated from various proteins and to this extent the information may be thrown into a semi-quantita-tive form of statement. But the precision of the information is very low indeed—we may not know the actual quantities present in the protein molecule within several hundred per cent in some cases.

I realize that this classification is one in which personal opinion plays a large part and it is given entirely with the object of eliciting discussion. There will probably be little debate about the hydroxy-acids, serine, threonine,* and hydroxyglutamic acid. Their presence (or absence) has been demonstrated in relatively few cases. It is pos-sible that serine may become better known through application of Rapoport's method[57] to convert it to glyceric acid which gives a color reaction with naphthoresorcinol, but this method awaits detailed study and full application.

---

* Shinn and Nicolet (Jour. Biol. Chem. **138**: 91. 1941) have recently described a method for the determination of threonine by means of oxidation with periodate which appears to be satisfactory, and mention that a method for serine is being developed.
[57] Rapoport, S. Biochem. Zeit. **281**: 30. 1935. **289**: 406. 1937.

A. Amino acids concerning which our information is little better than qualitative.

Serine	Isoleucine
Threonine	Norleucine
Hydroxyglutamic acid	Thyroxine
Valine	Diiodotyrosine

B. Amino acids for which methods of a considerable degree of probable accuracy have been proposed. These methods have been applied to very few proteins as yet.

Glycine	Proline
Alanine	Hydroxyproline
Leucine	Phenylalanine

C. Amino acids for which existing methods appear to give satisfactory results and which have been widely applied.

Cystine	Glutamic acid
Tyrosine	Arginine
Tryptophane	Histidine
Methionine	Lysine
Aspartic acid	

Isoleucine has seldom been demonstrated in protein hydrolysates, the custom having been to weigh the amino acids of the correct nitrogen or carbon content and report them as "leucines." With respect to norleucine, we have only a few statements, all of them merely qualitative.

Valine has been prepared in almost every case by the ester distillation method and subsequent fractional crystallization of the free acids and, until recently, this was for the most part true also for the leucines. This group of substances is separated from phenylalanine during distillation of the esters—this being in fact the chief purpose of the esterification procedure—and the separation of individual acids depends on the skill of the investigator in the use of fractional crystallization, of metallic salts (*e. g.* copper, lead, zinc), or of organic derivatives. These methods are, however, really preparation methods as distinguished from analytical methods and the yields are often notably poor. The inevitable errors introduced in the ester distillation procedure were discussed by Osborne[58] many years ago and are well recognized.

The inclusion of the iodine-containing amino acids thyroxine and diiodotyrosine in Group A will also be readily understood, in spite of

⁵⁸ Osborne, T. B., & Jones, D. B. Am. Jour. Physiol. 26:305. 1910.

the facts that Brand and Kassell[59] have recently developed new methods for them and that several other methods have been described. Although much labor has been expended on thyroxine isolation from thyroid gland tissue, and the yields obtained from glands of different sources are a matter of interest because of the physiological and pharmacological uses, there is very little information available on the exact thyroxine content of purified preparations of the protein thyroglobulin. For diiodotyrosine (iodogorgoic acid), which is also a component of thyroglobulin,[60] there is similarly little more than qualitative information.

The second group of amino acids in the table (Group B, TABLE 2) is debatable ground. Glycine, alanine, proline, oxyproline, leucine, and phenylalanine until recent years could be separated only by the Fischer procedure and, although much valuable data was secured, the results can only, even under the best conditions and with the most careful attention to details, have been approximations. There is no way in which one can calculate what the error in these determinations may have been. Doubtless many of the values in the literature approach the truth quite closely; others, however, must be much too low and some may even be too high.

These early data served to discriminate between various proteins in that they showed, for example, that prolamines were very low in glycine and high in proline, but they could have no possible significance for the calculation of molecular ratios between different amino acids nor for the calculation of minimal molecular weights of the protein.

In the past few years, the analytical chemistry of this group of amino acids has been placed upon a much more encouraging footing by the investigations of Bergmann. It is not too much to hope that his new methods, when generally applied to a wide range of purified proteins, may go far to provide the solution of the problem of the analytical chemistry of this group of substances.

The method for glycine[61] involves the precipitation of the complex compound of potassium trioxalato chromiate, $[Cr(C_2O_4)_3]K_3 + 3H_2O$, that is formed when the reagent is added, together with two volumes of alcohol, to the solution that contains glycine. Under carefully controlled conditions of temperature, acidity and time, a constant proportion of the glycine (about 88 per cent) is precipitated. With the use of this value, the quantities of glycine precipitated from a

[59] Brand, E. & Kassell, B. Jour. Biol. Chem. **131**: 489. 1939.
[60] Foster, G. L. Jour. Biol. Chem. **83**: 345 1929.
[61] Bergmann, M., & Fox, S. W. Jour. Biol. Chem. **109**: 317. 1935 Bergmann, M., & Niemann, C. Jour. Biol Chem. **122**: 577. 1937.

protein hydrolysate are corrected and an estimate is made of the glycine yielded.

The determination of alanine requires the previous removal of glycine and involves a series of small scale determinations of the nitrogen precipitated by dioxalatodipyridinochromiato acid[62] (dioxopyridic acid), $[Cr(C_2O_4)_2.(C_5H_5N)_2].H$ in order to find the proper quantity of reagent to add in order to precipitate the alanine. The weight of the salt is corrected for the small solubility to obtain the yield of alanine.

To determine proline, advantage is taken of the insolubility of the proline compound of tetrathiocyanato-dianilidochromiato acid, $[Cr(CNS)_4(C_6H_5.NH_2)_2].H$ (rhodanilic acid). Again a solubility correction must be ascertained and applied.

From the filtrate from the proline compound, hydroxyproline may be precipitated by the addition of Reinecke salt, $[Cr(CNS)_4(NH_3)_2].$ NH_4, together with pyridine. Purification of the hydroxyproline is difficult and the method is probably considerably less satisfactory as an analytical procedure than that employed for the other amino acids. Nevertheless higher yields, as compared with the results of other methods, were obtained.

A consideration of the difficulties and uncertainties involved in these precipitation methods—particularly with respect to the incompleteness of precipitation and consequent necessity for the use of solubility corrections—has led during the past two years to the development of methods of a new type in Bergmann's laboratory.[63] If to a solution of an amino acid a reagent is added in quantity *less* than sufficient to precipitate all the amino acid, a precipitate will form that contains, when equilibrium is reached and on the simplest assumption of binary salt formation, equimolecular quantities of the reagent and the amino acid. The mother liquor will contain a certain concentration of the ions of both reagent and amino acid. If a known amount of reagent has been employed, the concentration of the reagent left in the mother liquor can be calculated and an equation set up for the solubility product of the two ions. If a second experiment with a different amount of reagent is then carried out, and the solubility product calculated, these two results can be equated. In the resulting expression, the only unknown is the molar concentration of the amino acid in the solution at the beginning. This calculation assumes that molar con-

---

[62] Bergmann, M.   Jour. Biol. Chem. **122**: 569.   1938.

[63] Stein, W. H., Niemann, C., & Bergmann, M.   Jour. Am. Chem. Soc. **60**: 1703.   1938. Bergmann, M., & Stein, W. H.   Jour. Biol. Chem. **128**: 217.   1939.   Ing, H. R., & Bergmann, M.   Jour. Biol. Chem. **129**: 603.   1939.

centrations can be substituted for ionic concentrations, that activity coefficients do not change during the experiment, and, for success, requires rigid control of the physical conditions under which the experiment is conducted. The great advantage is that no knowledge of the solubility of the amino acid compound is required and, accordingly, the insidious errors are avoided that may creep in when the solubility of a compound in water is used as a correction for the solubility of the same compound in a solution of amino acids. The justification of the method is that, in spite of the assumptions involved, quantitative recoveries of amino acids, either in pure solution or in mixtures, can be readily secured.

New data have already been obtained in several cases by this method. For example, the glycine content of gelatin, previously given as 25.5 per cent, is now found to range from 26 to 27 per cent according to the source of the protein preparation. The proline, formerly 19.7 per cent, is now 17.5 per cent with a precision of about 2.5 per cent of this value. The revision downwards by the new method furnishes an illustration of the effect of errors in the earlier method.

In order conveniently to apply this method, Bergmann and Stein[64] have proposed an entirely new reagent for certain amino acids, namely naphthalene-β-sulfonic acid, the salts of which are referred to as nasylates, to distinguish them from the naphthylenesulfonyl derivatives. The selection of this substance, which yields salts that are only moderately insoluble, is possible because a definite and appreciable solubility is a distinct advantage in the application of the new solubility product method and simple binary salt formation with the amino acid is desirable. With this reagent, the determination of leucine, arginine, and probably other amino acids, is possible. The salts of isoleucine and valine are more soluble than those of leucine and, accordingly, these amino acids do not interfere with leucine determinations in proteins. The determination of phenylalanine with this new reagent or some substance of analogous properties also appears to be a definite possibility.

As yet this method has not been applied to more than a few proteins. But the accuracy with which solutions of amino acids, or simple mixtures of amino acids can be analyzed suggests that most valuable results are to be anticipated.

Dr. Bergmann kindly permits me to say at this point that an extensive series of sulfonic acid reagents is being investigated and that suitable reagents for many specific precipitations have been found.

⁶⁴ **Bergmann, M., & Stein, W. H.** Jour. Biol. Chem. **129**: 609. 1939.

The third group of amino acids (Group C, TABLE 2), which contains those substances for which analytical methods of some accuracy exist, also furnishes material for considerable debate. However one may answer the question, what is the best method to determine cystine, one is certain to encounter argument, and since this problem, and the allied problems that concern methionine, tyrosine, and tryptophane have been discussed at great length in the literature of the past few years, it seems scarcely necessary to go into it at all deeply in this place. One is perhaps justified in setting up as a criterion for judgment among the various methods the requirement that they should yield substantially the same result when careful comparative tests are conducted. Hess and Sullivan[65] have recently determined the cystine content of several proteins by five different current methods, those of Sullivan, of Okuda, of Shinohara, of Folin and Marenzi, and of Vickery and White. Comparison of the results shows a remarkably close agreement for four of the five methods, only that of Folin and Marenzi being significantly different. Such results give a considerable degree of confidence in the accuracy of each of these four methods and suggest that choice among them can properly be made on grounds of convenience or personal preference. Other similar comparisons have been made in Chibnall's laboratory by Bailey.[66]

The two dicarboxylic amino acids have received far less attention than most of the others. The methods employed have been gradually developed from the early purely preparation type of procedure and have recently been considerably refined in Chibnall's laboratory. There is still room for much improvement, however, and it is doubtful if most of the data in the literature are as satisfactory as the results for cystine, methionine, tyrosine, and tryptophane.

Of the three basic amino acids, arginine can now be determined with great accuracy. There are a number of methods available, each with advantages for certain purposes. A careful comparison of recent results in my laboratory of a new modification of the flavianic acid method[67] with published values by other methods suggests a definite superiority for the new method on grounds both of accuracy and of precision. Nevertheless the agreement of several of the other methods, notably the beautiful enzymatic method of Hunter, with the results of the new procedure indicates that there is latitude for choice.

Histidine and particularly lysine, however, are in a less favorable

65 Hess, W. C., & Sullivan, M. X. Jour. Biol. Chem. 128: 93. 1939.
66 Bailey, K.   Biochem. Jour. 31: 1396. 1937.
67 Vickery, H. B.   Jour. Biol. Chem. 132: 325. 1940.

position. There are several procedures whereby these substances may be determined and, a year ago, I should not have hesitated to claim that the large-scale silver precipitation method, originally devised by Kossel and more recently modified in my laboratory, yields results that are highly trustworthy. Today I am not so certain. By the use of the new flavianic acid procedure, appreciably higher results for arginine have been secured in many cases than we were able to get by the silver precipitation method and it is only natural to inquire whether renewed attention to specific methods for histidine and lysine may not likewise result in a marked improvement. For the present, then, I can only say that the results of the large-scale silver precipitation method are probably the most trustworthy available, but I doubt if they can be regarded as completely satisfactory until new methods have been devised and the data have been confirmed.

To sum up what may be regarded as a conclusion with respect to the trustworthiness of analytical data on proteins, I think that we have some reason to be content with many of the figures for cystine, tyrosine, tryptophane, methionine, and arginine. There are a few values for glutamic and aspartic acid worthy of serious consideration and some of the data for histidine and lysine are probably quite close to the truth. Doubtless, when Bergmann's new methods have been more fully developed, we shall have valuable data for several of the mono-amino acids, but as yet we are awaiting general application of these methods to proteins that can be regarded as homogeneous chemical entities.

This conclusion may seem unduly pessimistic to some, but I sincerely feel that truly accurate values for a few amino acids in a protein that has been shown by physico-chemical methods to be a substance rather than a mixture are of greater value for our understanding of the complexities of the protein molecule than is a mass of approximations.

APPLICATION OF ANALYTICAL RESULTS

Analytical data for the amino acids derived from proteins may be roughly assigned to two categories. There is a relatively large group of determinations obtained by Fischer, Osborne, and Abderhalden and their respective associates, and by Jones, Dakin, and a few subsequent workers in which the object was to account for as large a part of the protein molecule as possible. Compilations of

data for many proteins can be made today which add up to a very satisfactory total and, in the cases of zein, gelatin, and a few others, this total accounts for 90 per cent or more of the molecule.   This information is of great value even though we know that many of the individual determinations, such as those of glycine, alanine, valine, leucine, etc., do not have a high degree of probable accuracy. Numerous attempts have been made to calculate molecular ratios from such data, with a moderate degree of success, but there is still so much doubt about the accuracy of the details that the results can be regarded only as approximations.

In another category, however, are the results of the determinations of the amino acids in Group C in TABLE 2. For most of these substances the methods permit a series of check determinations without the expenditure of too much labor and an estimate of both accuracy and precision becomes possible.   Accordingly the data can properly be employed to calculate molecular ratios and minimal molecular weights.

In the following tables* (TABLES 3 to 7) a number of these calculations are presented.   The data that have been selected are taken from studies that show evidence of unusual care and attention to controls.   In most cases the values are averages of several or of many separate determinations and, in their selection, consideration has been given to the results of other investigators who have obtained values of a closely agreeing magnitude.   It is not suggested that these are the *best* values, that being largely a matter of opinion, nor are they invariably the highest values; they are simply the results of unusually careful analytical studies.

Many methods of presenting such data are possible but, in the following tables, since the emphasis is to be placed upon the analytical aspect of the problem, the first five columns give the calculations of the minimal molecular weight of the protein derived from each item.   This minimal molecular weight is divided into the value assigned to the true molecular weight, obtained from physico-chemical measurements, and the nearest integer to the quotient is given in Column 6.   Multiplication of the figures in Columns 5 and 6 gives the calculated molecular weight in Column 7, and the result is calculated as a percentage of what may be called the "physical" molecular weight in Column 8.   From this figure the agreement of the calculated "chemical" molecular weight may be easily appreciated.

TABLE 3

HEMOGLOBIN (HORSE)

Assumed Molecular Weight 66700

	Observed per cent	Author- ity	Moles per gm. X 10⁻⁵	Minimal molecular weight	Assumed number of atoms or residues	Cal- culated molecular weight	Percentage of 66700
1	2	3	4	5	6	7	8
Iron	0.335	1	5.99	16660	4	66640	99.9
Sulfur	0.390	1	12.17	8220	8	65760	98.6
Sulfide sulfur	0.191	2	5.96	16780	4	67120	100.7
Sulfur	0.57	3	17.8	5623	12	67480	101.2
Arginine	3.59	4	20.62	4849	14	67890	101.8
Histidine	7.64	5	49.26	2030	33	66990	100.4
Lysine	8.10	5	55.44	1804	37	66750	100.1
Tyrosine	3.15	6	17.4	5749	12	68980	103.4
Tryptophane	1.28	6	6.27	15940	4	63760	95.6
Cystine	0.41	7	1.71	58590	1	58590	88
Aspartic acid	8.9	8	66.9	1496	45	67320	100.9
Glutamic acid	6 3	8	42.8	2335	29	67710	101.5
Amide N	1.01	8	72.1	1387	48	66580	99.8

If tryptophane is 1.22%, millimols per gm. = 5.98; min. molecular weight = 16730; for 4 mols 66920

If cystine is 0.36%, millimols per gm. = 1.49; min. molecular weight = 66730; for 1 mol 66730

If cystine is 0.72%, millimols per gm. = 2.99; min. molecular weight = 33370; for 2 mols 66740

[1] Zinoffsky, O. Zeit. physiol. Chem. 10: 16. 1886.
[2] Schulz, F. Zeit. physiol. Chem. 25: 16. 1898.
[3] Valer, J. Biochem. Zeit. 190: 444. 1927.
[4] Vickery, H. B. Jour. Biol. Chem. 132: 325. 1940.
[5] Vickery, H. B., & Leavenworth, C. S. Jour. Biol. Chem. 79: 377. 1928.
[6] Folin, O., & Marenzi, A. D. Jour. Biol. Chem. 83: 89. 1929.
[7] Vickery, H. B., & White, A. Jour. Biol. Chem. 99: 701. 1933.
[8] Chibnall, A. C., & Bailey, K. Personal communication.

It is obvious that when the assumed number of residues or atoms in Column 6 is small, the probable validity of the calculated "chemical" molecular weight is high. In this case the accuracy of the analytical value receives its sharpest test and at the bottom of TABLE 3 a calculation is given which shows how important the second place of decimals in the tryptophane and cystine values may be.

On the other hand, when the number in Column 6 is large, there is a considerable element of uncertainty in the validity of the agreement between the "chemical" and the "physical" molecular weights. Frequently a neighboring integer may equally well be chosen.

If it were possible to assign weighting factors to these several analytical results, one might calculate a molecular weight that would express the sum total of our best knowledge derived from analytical chemistry. At the moment, however, there is no objective way in

which this can be done, although it is obvious that some of the results are more highly significant than others.

Hemoglobin

The iron value of Zinoffsky has been repeatedly confirmed, notably by Valer in recent years, and it is from this value that the assumed molecular weight is calculated (TABLE 3). The physicochemical results of Adair and of Svedberg are in very close agreement.

Zinoffsky's sulfur value differs from that of Valer in a most interesting way. The analytical methods employed were widely different; Zinoffsky used an oxidation procedure while Valer employed the reduction method of ter Meulen in which the protein is heated in an atmosphere of hydrogen and the gases are passed over a platinum catalyst, sulfur being determined as hydrogen sulfide by iodine titration. Valer's values were obtained on the hemoglobin of individual animals and, with the dog, he found evidences of definite differences that he suggested might have a genetic origin. Whether or not this is the explanation of the results for the horse is a problem that awaits study. From the analytical standpoint it is important to note that *both* values fit well with the rest of the data.

The chemical meaning of the sulfide sulfur determination is also still unexplained. The value of Schulz given in the table is identical with Osborne's and it is significant that, in this case, it is twice as great as the cystine sulfur. In other cases (TABLE 7), sulfide sulfur agrees with the cystine sulfur and the result in the present case suggests that the Vickery and White cystine value may be considerably too low.

Stern, Beach, and Macy[68] have recently studied the globins of several species, cystine being determined by the polarographic method of Brdicka. The result for horse globin was 0.77 per cent, and they report a value of 0.81 per cent by Graff, Maculla, and Graff's[69] modification of the Vickery and White method. On the assumption that horse hemoglobin contains 2 per cent of haem, the polarographic value for cystine would be 0.75 per cent and the calculations at the bottom of TABLE 3 show that this agrees well with the assumption that horse hemoglobin contains 2 moles of cystine per molecule. Furthermore such an assumption would bring the value for sulfide sulfur in line with observations on other proteins. Uncertainty still exists, however, with respect to the total sulfur content of horse hemoglobin,

[68] Stern, A., Beach, E. F., & Macy, I. G.   Jour. Biol. Chem. **130**: 733.   1939.
[69] Graff, S., Maculla, E., & Graff, A. M.   Jour. Biol. Chem. **121**: 81.   1937.

TABLE 4

EGG ALBUMIN (HEN)

Assumed Molecular Weight 36000 and 40000

	Observed per cent	Authority	Moles per gm. $\times 10^{-5}$	Minimal molecular weight	Assumed molecular weight 36000			Assumed molecular weight 40000		
					Assumed no. of atoms or residues	Calculated molecular weight	Percentage of 36000	Assumed no. of atoms or residues	Calculated molecular weight	Percentage of 40000
1	2	3	4	5	6	7	8	9	10	11
Arginine	5.66	1	32.52	3075	12	36900	102.5	13	39970	99.9
Histidine	1.48	2	9.766	10240	4	40960	113.9	4	40960	102.3
Lysine	4.97	2	34.02	2940	12	36240	100.7	14	41160	102.9
Tyrosine	3.97	3	21.93	4561	8	36490	101.4	9	41050	102.6
Tryptophane	1.18	3	5.78	17300	2	34600	96.1	2	34600	84.5
Cystine	1.78	4	7.41	13500	3	40500	112.6	3	40500	101.3
Cysteine	1.78	4	14.7	6803	5	34020	94.5	6	40820	102.1
Methionine	5.23	4	35.05	2853	13	37090	103.0	14	39940	99.8
Glutamic acid	13.96	5	95.13	1052	34	35760	99.3	38	39980	99.9
Aspartic acid	6.07	5	45.60	2761	13	35890	99.6	15	41420	103.5
Sulfur	1.616	6	50.43	1983	18	35690	99.1	20	39660	99.1
Phosphorus	0.097	6	3.13	32010	1	32010	88.8	1	32010	80.0
Sulfide sulfur	0.491	6	15.32	6528	6	39170	108.8	6	39170	97.9
Amide Nitrogen	0.97	7	69.3	1443	25	36070	100.2	28	40390	100.9

[1] Vickery, H. B. Jour. Biol. Chem. 132: 325. 1940.

[2] Vickery, H. B., & Shore, A. Biochem. Jour. 26: 1101. 1932.

[3] Folin, O., & Marenzi, A. D. Jour. Biol. Chem. 83: 89. 1929.

[4] Kassell, B., & Brand, E. Jour. Biol. Chem. 125: 145. 1938.

[5] Calvery, H. O. Jour. Biol. Chem. 94: 613. 1931.

[6] Osborne, T. B. Jour. Am. Chem. Soc. 24: 140. 1902.

[7] Shore, A., Wilson, H., & Stueck, G. Jour. Biol. Chem. 112: 407. 1935.

as well as the cystine content and it is obvious that further study of the sulfur and the sulfur-containing amino acids of hemoglobin obtained from a known source is needed.

The values for the bases and the dicarboxylic acids require little comment. Of these the arginine is known with the greatest certainty and the proportion present is such that 14 may be selected for the number of residues with some confidence; neither 13 nor 15 fits the data at all well. For the others the choice is not so clearly defined. The chief significance of the amide nitrogen values is to show that both aspartic and glutamic acids must be present in part as their respective amides.

Egg Albumin

The calculations for egg albumin (TABLE 4) introduce a further complexity inasmuch as the molecular weight derived from physicochemical considerations is still uncertain. The value 36000 is indicated by osmotic pressure and by the early ultracentrifuge data; later figures of Svedberg indicate a value near 40000. Calculations have accordingly been made on both bases and, in general, the agreement of the analytical data is better when the higher molecular weight is chosen. Furthermore the higher value brings the cystine and methionine values of Kassell and Brand in line with Osborne's very carefully determined total sulfur, and the cystine also agrees with the sulfide sulfur, although whether or not this is evidence is still debatable. Calvery's phosphorus does not lead to a decision and obviously the phosphorus content as well as the cystine and tryptophane would repay further study.

A recent paper of Bernhart[70] gives new values for tyrosine (3.89 per cent) and tryptophane (1.13 per cent) which agree closely with the earlier values of Folin and Marenzi used in TABLE 4. In addition there is a new value for phenylalanine by the Kapeller-Adler method of 5.37 per cent. This value, if the method is assumed to be accurate, leads to a minimal molecular weight of 3147 and accordingly to estimates of a molecular weight of 37700 for 12 units and 40900 for 13. Bernhart's analysis of his own data and that of others led him to an estimate of 36900 as a probable molecular weight of egg albumin.

Insulin

The situation (TABLE 5) with respect to the zinc content of crystalline insulin is similar to that of the sulfur content of hemoglobin.

---

[70] Bernhart, F. W.   Jour. Biol. Chem. **132**: 189.   1940.

TABLE 5

INSULIN

Assumed Molecular Weight 35100[7]

	Observed per cent	Author-ity	Moles per gm. $\times 10^{-5}$	Minimal molecular weight	Assumed number of atoms or residues	Cal-culated molecular weight	Percentage of 35100
Zinc	0.52	1	7.95	12570	3	37710	107.4
Zinc	0.36	2	5.51	18160	2	36320	103.5
Sulfur	3.34	3,4	104.2	960	36	34560	98.5
Cystine	12.5	3	52.0	1922	18	34590	98.5
Cysteine	12.6	3	104.1	961	36	34590	98.5
Tyrosine	12.2	5	67.4	148.5	24	35640	101.5
Amide N	1.55	6	110.8	903	39	35220	100.3
Cadmium*	0.77	1	6.85	14600	2	29200	81.3
Cobalt**	0.44	1	7.47	13400	3	40200	114.5

* If Scott and Fisher's lowest cadium value of 0.72 per cent is taken, the minimal molecular weight is 15610 and the calculated molecular weight for 2 atoms is 88.9 per cent of 35100.

** If Scott and Fisher's highest cobalt value of 0.46 per cent is taken, the minimal molecular weight is 12810, and the calculated molecular weight for 3 atoms is 109.4 per cent of 35100. Their lowest value of 0.41 per cent calculated for 2 atoms gives 81.8 per cent.

1 Scott, D. A., & Fisher, A. D. Biochem. Jour. 29: 1048. 1935.
2 Cohn, E. J., Ferry, J. D., Livingood, J. J., & Blanchard, M. H. Science 90: 183. 1939.
3 Miller, G. L., & du Vigneaud, V. Jour. Biol. Chem. 118: 101. 1937.
4 du Vigneaud, V., Miller, G. L., & Rodden, C. J. Jour. Biol. Chem. 131: 631. 1939.
5 du Vigneaud, V., Jensen, H., & Wintersteiner, O. Jour. Pharmacol. 32: 367. 1928.
6 du Vigneaud, V. Cold Spring Harbor Symposia Quant. Biol. 6: 275. 1938.
7 Svedberg, T. Nature 127: 438. 1931.

Scott and Fisher's value fits well for 3 zinc atoms, that of Cohn and his associates fits even better for 2 atoms. Scott and Fisher's cadmium and cobalt values in crystals obtained with these metals do not resolve the problem. However, the sulfur, cystine, and tyrosine values fit the requirements of the ultracentrifuge determinations of Svedberg very satisfactorily.

Serum Albumin

Serum albumin (TABLE 6) presents greater difficulties for discussion from the analytical point of view than the proteins hitherto mentioned since there is every reason to believe, today, that what has usually been isolated as serum albumin is a mixture of proteins that differ from each other, for example, in tryptophane content. Early data on the tryptophane of serum albumin indicated 0.53 per cent.[71] This result was obtained from a sample crystallized twice by Wyman in Cohn's laboratory. Another sample prepared by Ferry gave 0.52 per cent of tryptophane. Hewitt's more recent work on serum albumin[72] indicates that fractions can be secured of which the "purest"

71 Folin, O., & Marenzi, A. D. Jour. Biol. Chem. 83: 89. 1929.
72 Hewitt, L. F. Biochem. Jour. 30: 2229. 1936.

yielded 0.26 and 0.30 per cent of tryptophane, while the albumin separated from the mother liquors gave as much as 1.0 per cent.   On the other hand, the most recent work of McMeekin in Cohn's labora-tory[78] has resulted in the crystallization of a carbohydrate-free serum albumin fraction of nitrogen content 16.2 per cent in the form of a sulfate and 16.8 per cent in the isoelectric condition.   Hewitt's material had only 14.1 to 14.4 per cent of nitrogen.   McMeekin's preparation yielded 0.53 per cent of tryptophane, by Lugg's modification of the Folin and Marenzi procedure, and thus resembles the earlier prepara-tions of Wyman and Ferry from the same laboratory.   Furthermore it was demonstrated to be essentially homogeneous in the Tiselius ap-paratus and to have a constant solubility.   This provides a second

TABLE 6

SERUM ALBUMIN (HORSE)

Assumed Molecular Weight 73000

	Observed per cent	Author-ity	Moles per gm. $\times 10^{-5}$	Minimal molecular weight	Assumed number of atoms or residues	Cal-culated molecular weight	Percentage of 73000
Tryptophane	0.26	1	1.275	12440	6	74640	102.2
Tryptophane	0.53	2	2.597	38500	2	77000	105.4
Tyrosine	4.79	1	26.4	3781	19	71840	98.4
Amino N	1.07	1	76.4	1308	56	73250	100.3
Cystine	6.04	3	25.1	3977	18	71590	98.0

1 Hewitt, L. F. Biochem. Jour. 30: 2229. 1936.
2 McMeekin, T. L. Jour. Am. Chem. Soc. 61: 2884. 1939.
3 Folin, O., & Marenzi, A. D. Jour. Biol. Chem. 83: 103. 1929

case of a protein of animal origin in which two widely different values for the same component seem equally well supported; the discrepancy between the two sulfur values of horse hemoglobin has already been mentioned.   Valer has pointed out the possibility that the genetic origin of the hemoglobin of the dog may have significance in its effect upon the sulfur content.   One may well ask if this may be a factor in the cases of the hemoglobin and the serum albumins of the horse.

There is a greater degree of uniformity in the tyrosine values for serum albumin, the early ones of Folin and Marenzi being 4.66 and 4.77 per cent respectively.   Hewitt's purest fractions gave 4.74 and 4.79 per cent as compared with earlier less carefully fractionated material of 4.7 per cent.[74]   The more soluble fractions, however, gave 5.38 and 6.06 per cent.

73 McMeekin, T. L. Jour. Am. Chem. Soc. 61: 2884. 1939.
74 Hewitt, L. F. Biochem. Jour. 28: 2080. 1934.

TABLE 7

EDESTIN (HEMPSEED)

Assumed Minimal Molecular Weight 56000

Molecular Weight 304000[9]

	Observed per cent	Author-ity	Moles per gm. $\times 10^{-5}$	Minimal molecular weight	Assumed number of atoms or residues	Cal-culated molecular weight	Percentage of 56000
Arginine	16.76	1	96.29	1039	54	56100	100.2
Histidine	2.08	2	13.42	7456	8	59650	106.5
Lysine	2.19	2	14.98	6673	8	53380	95.3
Tyrosine	4.54	2	25.07	3988	14	55830	99.7
Tryptophane	1.46	3	7.16	13970	4	55880	99.8
Cystine	1.36	4	5.66	17660	3	52980	94.6
Cysteine	1.36	4	11.23	8904	6	53420	95.3
Methionine	2.39	5	16.02	6243	9	56190	100.3
Glutamic acid	19.2	6	130.5	766	73	55920	99.8
Aspartic acid	10.2	6	76.6	1305	43	56120	100.2
Sulfur	0.880	7	27.46	3642	15	54630	97.5
Sulfide sulfur	0.346	7	10.79	9264	6	55580	99.2
Amide N	1.88	8	134.3	7450	75	55870	100.0

1 Vickery, H. B. Jour. Biol. Chem. 132: 325. 1940.

2 Vickery, H. B., & Leavenworth, C. S., Jour. Biol. Chem. 76: 707. 1928.

3 Folin, O., & Marenzi, A. D. Jour. Biol. Chem. 83: 89. 1929.

4 Bailey, K. Biochem. Jour. 31: 1396. 1937.

5 Baernstein, H. D. Jour. Biol. Chem. 106: 451. 1934.

6 Jones, D. B., & Moeller, O. Jour. Biol. Chem. 79: 429. 1928.

7 Osborne, T. B. Jour. Am. Chem. Soc. 24: 140. 1902.

8 Osborne, T. B., & Harris, I. F. Jour. Am. Chem. Soc. 25: 323. 1903.

9 Svedberg, T. Ind. Eng. Chem., Anal. Ed. 10: 113. 1938.

Folin and Marenzi carried out cystine determinations on the crystalline samples already mentioned finding 6.02 and 6.06 per cent respectively. Unfortunately no data for sulfur are given by Hewitt, nor by McMeekin. In a later paper, however, Hewitt[75] states that the cystine content of his crystalbumin is 5.8 per cent which agrees well with the Folin and Marenzi value.

Edestin

The calculations on edestin (TABLE 7) are introduced largely to show how the analytical approach to the determination of molecular weight, or rather to the selection of a most probable molecular weight, breaks down when a protein of high molecular weight is taken. The best that can be done is to assume an arbitrary value that conforms closely with some of the data and to develop the consequences of the choice. In the present case, the tryptophane value is taken as correct and it is assumed that the molecular weight is a multiple of 14000.

75 Hewitt, L. F. Biochem. Jour. 31: 360. 1937.

In order to accommodate the cystine value, 56000 must be selected and the table is constructed on the assumption that this is a sub-multiple of the true molecular weight. Five times this is 280000, six times is 336000, and neither of these agrees very well with Sved-berg's ultracentrifuge result.

The selection of analytical data is fairly easily made for this protein since it has repeatedly been carefully studied. The tyrosine value, 4.54 per cent, was obtained by Folin and Marenzi on a sample recorded as "not crystallized." This, however, does not necessarily connote a lack of purity. A crystalline sample obtained from Osborne gave 4.28 per cent tyrosine. Other values in the literature, however, are 4.58 per cent[76] and 4.53 per cent[77] and the weight of evidence therefore favors the higher result of Folin and Marenzi. These same workers have also obtained tryptophane values that agree well. Folin and Marenzi got 1.45 and 1.46 per cent on their two prepara-tions, Looney got 1.52 per cent, and Folin and Ciocalteu 1.51 per cent.

The cystine value of 1.36 per cent obtained by Bailey, in the course of a careful comparison of the results of various cystine methods on this protein, is supported by Folin and Marenzi's value of 1.35 per cent, Sullivan and Hess' (iodometric) value of 1.32 per cent, and is only a little higher than Vickery and White's value of 1.25 per cent by the cuprous mercaptide method.

The methionine value of Baernstein, obtained by the titration of homocysteine thiolactone, was confirmed almost precisely (2.38 per cent) in a later paper in which the volatile iodide method was used. Bailey has also obtained a value of 2.3 per cent by the volatile iodide method. None of these values contains the corrections estimated from recovery experiments by Kassell and Brand.

The sulfur value of Osborne falls in line with these results for cystine and methionine and his sulfide sulfur again agrees quite well with the cystine. The amide nitrogen value indicates that all of the glutamic acid may be present as glutamine in this protein and all of the aspartic acid may be free. That this is not the case is shown by Damodaran's isolation of asparagine from edestin.[55]

SUMMARY

The view that amino acids are the proximate components of the protein molecule was first clearly expressed by Ritthausen. The implications of this idea were explored by Kossel, Fischer, and Hof-

[76] Looney, J. M.   Jour. Biol. Chem. **69**: 519.   1926.
[77] Folin, O., & Ciocalteu, V.   Jour. Biol. Chem. **73**: 627.   1927.

meister but were largely ignored by Kühne and his pupils. During the past one hundred and twenty years, more than forty different substances or preparations have been isolated from proteins and claimed, with more or less evidence, to represent individual amino acid units of which the protein molecule is formed. These are listed in TABLE 1 and are classified into four groups. Eighteen of them are universally recognized as common protein constituents, and seven more occupy a somewhat special position, either by reason of their limited distribution or because certain details of their mode of combination in the protein are not understood.

In addition there are five known amino acids, found in plant tissues and, in a few cases, in animal tissues which are by no means improbable, though doubtless rare, constituents of proteins although this has not yet been shown. Some seventeen other substances have been at one time or another reported as protein constituents but without wholly satisfactory evidence.

The well known protein amino acids are classified in TABLE 2 in accordance with our knowledge of their analytical chemistry. For eight amino acids there are at present no dependable analytical methods whatever; knowledge of most of them rests entirely on relatively crude isolation methods. For an additional group of six mono-amino acids, fairly good preparation methods exist, but these are far from quantitative. The current researches of Bergmann are, however, revealing methods by which these substances can be determined with a high degree of probable accuracy and satisfactory determinations of them in a series of proteins are to be looked for in the near future.

A final group can be formed of those amino acids for which colorimetric or isolation methods of high probable accuracy have been developed. It is with these substances alone that trustworthy molecular ratios and minimal molecular weights can be calculated at the present time.

In TABLES 3 to 7 are presented calculations of the amino acid composition of five proteins which have been shown to be nearly, if not entirely, homogeneous chemical substances. Selections of data for these calculations have been made from papers which contain evidence that unusual care has been exercised in the effort to secure accurate results. In many cases the data selected have been confirmed by other methods or by other investigators. The results of the calculations of the minimal molecular weights, in combination with molecular weights derived from physico-chemical considera-

tions, permit the selection of most probable integers for the number of amino acid residues per molecule. A picture of the composition of the protein molecule can thus be drawn, and from the agreement, or lack of agreement, of the molecular weights calculated from the analytical data with the value secured by physical methods, an estimate of the trustworthiness of both chemical and physical data can be made. It is found that a surprising number of analytical determinations fit well with each other and provide valuable cumulative evidence of the validity of the assumptions upon which these calculations rest.

Full advantage is also taken of the cases, such as hemoglobin and insulin, where a metal atom forms one of the units of the structure, and the relationships of the two known sulfur-containing amino acids with the total sulfur are also shown to be satisfactory in most cases.

EVIDENCE FROM PHYSICAL CHEMISTRY REGARDING THE SIZE AND SHAPE OF PROTEIN MOLECULES FROM ULTRA-CENTRIFUGATION, DIFFUSION, VISCOSITY, DIELECTRIC DISPERSION, AND DOUBLE REFRACTION OF FLOW

By J. L. Oncley

*From the Department of Physical Chemistry, Harvard Medical School,
Boston, Massachusetts*

The estimation of size and shape of protein molecules has been attempted by many kinds of physico-chemical measurements. Nearly all of these methods have both experimental and theoretical uncertainties. In the attempt to extend the data so obtained to what can well be called the second approximation, some of the uncertainties have perhaps been forgotten in our enthusiasm. In this paper we shall try to discuss the major uncertainties, but shall avoid any unnecessary discussion of experimental techniques. We shall further confine our discussion as much as possible to cases where only one protein component is present, neglecting the interactions of one protein on another, and the problems involved in the dissociation or association of protein molecules.

The very fact that these second approximations are under discussion here is perhaps the highest tribute that we can pay to Professor The Svedberg and his associates, to whom we are largely indebted for many techniques[1] which were in the form of preliminary investigations only fifteen years ago.

SEDIMENTATION EQUILIBRIUM

From a theoretical standpoint, the sedimentation equilibrium method of Svedberg is perhaps the most desirable method for the evaluation of molecular size.[2] It involves the measurement of the protein concentration, c, as a function of the distance from the center of rotation, x, when an equilibrium between centrifugal and diffusion

[1] Svedberg, T., & Pedersen, K. O. "The Ultracentrifuge". Oxford Univ. Press London. 1940.

[2] Svedberg, T. Zeit. physik. Chem. 121:65. 1926. Tiselius, A. Zeit. physik. Chem. 124:449. 1926. Pedersen, K. O. Zeit. physik. Chem. A170:41. 1934. See reference 1, pp. 48–57.

forces is set up.   The relation between these quantities, the speed
of rotation, ω, the partial specific volume, $\bar{v}$, and the molecular weight,
M_e, can be shown thermodynamically to be independent of the shape
and hydration of the protein, and given by the equation

$$M_e = \frac{2\,RT}{(1 - \bar{v}\rho)} \cdot \frac{\ln c_2/c_1}{\omega^2(x_2{}^2 - x_1{}^2)} \tag{1}$$

Here R is the gas content $(8.313 \times 10^7$ ergs per mole degree C.$)$, T,
the absolute temperature, and ρ the density of the solution.   This
method has been used by Svedberg and his co-workers for the deter-
mination of the molecular weights of most of the substances which
we intend to discuss in this symposium, and it is difficult for us
to evaluate the probable accuracy of the determinations.   A discus-
sion of the possible sources of error, however, will be of some use in
weighing evidence from this source.

Probably the most obvious source of error is the possibility of not
having set up a true equilibrium between the centrifugal and dif-
fusion forces.   For molecules of the size of those reported here this
requires times ranging from several days to perhaps a week.[3]   This
is usually tested by taking a series of pictures at 12- or 24-hour in-
tervals, until perhaps three consecutive pictures give identical results
within the expected experimental error.   The possibility exists, how-
ever, that this state may not be one of true equilibrium, since the
additional effects of thermal gradients and vibration may be affect-
ing the results.   Svedberg and the other workers in this field have
taken great pains to guard against this possibility.

Another possible source of error involves the possibility of some
change taking place in the protein molecules because of the long
times involved in these measurements.   The necessity for the material
to remain at 20° or 25° C. for prolonged periods without any change
restricts this method to only the most stable of the proteins.

The other sources of error are the limiting accuracies with which
the concentrations can be measured in the ultracentrifuge, the con-
stancy of the speed of the rotor, and the values used for the partial
specific volume of the protein, .   We will reserve the discussion of
this last point till later.   The error in evaluating the concentration
as a function of the distance is difficult to ascertain.   It would
obviously depend a great deal upon the method used; those in common

[3] Mason, M., & Weaver, W. Phys. Rev. **23**: 412. 1924. Weaver, W. Phys. Rev.
27: 499. 1926; Zeit. physikal. Chem. **43**: 396. 1927; **49**: 311. 1928. Archibald, W. J.
Phys. Rev. **53**: 746. 1938; **54**: 371. 1938. See reference 1, p. 56–57.

use are the scale-displacement method of Lamm, the slit method of Toepler, and the light adsorption method of Svedberg. The first of these is probably the most accurate for this purpose, and the most widely used at present. The experimental errors involved in comparing and calculating concentrations from the scale-displacement method are usually of a slightly larger magnitude than those to be discussed in the study of diffusion constants, and could readily amount to perhaps 5%.[4] Measurements are usually recorded by merely giving values of the molecular weight as calculated in different portions of the cell, values of x_1 and x_2 being chosen at 0.5 mm. intervals. The parts of the cell nearer the center, where c changes very slowly, are poorly suited to this means of calculation and are often omitted. The practice of plotting log c against x^2 would seem a more fitting one,[5] the curve so obtained being linear if the solution is monodisperse. The slope would be equal to

$$M_s\omega^2(1 - \bar{v}\rho)/4.606 \; RT.$$

When the system under analysis is polydisperse, the plot of log c against x^2 is not linear.[6] It has been shown that the values for the molecular weight calculated at various values of x in such a system increase as x increased. The absolute values obtained depend to a large extent upon the method used for the calculation. Three different ways of averaging the weight at any values of x are in use: (1) the number-average $(M_n = \Sigma_i(n_iM_i)/\Sigma_i n_i)$; (2) the weight-average $M_w = \Sigma_i(n_iM_i^2)/\Sigma_i(n_iM_i) = \Sigma_i(w_iM_i)$; (3) the z-average $M_z = \Sigma_i(n_iM_i^3)/\Sigma_i(n_iM_i^2)$. Here n_i and w_i represent the number and total weight of molecules of molecular weight M_i. Lansing and Kraemer[7] have discussed this problem very thoroughly, and give equations for the calculation of any of these quantities. The z-average is most easily obtained if the Lamm scale method is used, but weight-averages

---

[4] Lansing and Kraemer state[7] that the errors, "rarely exceed 10%, except in extremely unfavorable cases." Pedersen[1] (p. 305) states that "the accuracy of M determined from a *single* sedimentation equilibrium experiment cannot be put higher than $M\pm$ (5–10%)."

[5] When values of M are calculated from equation (1) using equal $x_2 - x_1$ intervals, it can be rather easily shown that all but the initial and final concentrations cancel because of the logarithmic nature of the equation. (See Roseveare, W. E. Jour. Am. Chem. Soc. **53**: 1651. 1931). This is mentioned by Pedersen[1] (p. 305) when he says "when only the average M is wanted, it is sufficient to know the concentration at the top and the bottom of the cell, since all the intermediate values cancel out when the mean is taken. This is, however, a drawback, as just the concentrations at the top and at the bottom of the cell are the ones that are most uncertain, since these can only be found by extrapolation."

[6] The polydispersity of a protein solution can be detected more readily by observations of the sedimentation velocity of the solution, as is discussed in a later section.

[7] Lansing, W. D., & Kraemer, E. O. Jour. Am. Chem. Soc., **57**: 1369. 1935. See also reference 1 pp. 342–9.

can be obtained by means of a graphical integration of the z vs. x curve. Number-averages cannot, in general, be determined with the reliability of the other averages, but are extremely useful since it is this number-average which is to be compared with osmotic pressure methods. The calculation of a "non-uniformity coefficient" can be obtained from any of these averaging methods. This problem will not be further discussed here since all of the molecules under discussion here have been shown to be monodisperse.

In some cases errors are introduced because of deviation of the protein from Henry's law. By keeping the concentration of protein at low values it is found that this source of error would apply only in extreme cases, and probably does not introduce any appreciable error for the molecules under discussion. Very large degrees of asymmetry will cause deviation from Henry's law to occur at lower protein concentration.

OSMOTIC PRESSURE

Another method which can be derived along strictly thermodynamic lines[8] is through the measurement of the osmotic pressure of solutions under proper conditions. It has been shown by Sörensen, Adair, Burk and others[9] that such measurements of osmotic pressure can lead to reliable molecular weights when proper corrections for membrane effects are made, and when the observations are made in sufficiently dilute protein concentrations.

In order to obtain such molecular weights, we must take account of (1) the "ion-difference pressure," p_i, (briefly described as the correction term for the unequal distribution of diffusible ions); (2) the effect of forces between ions and molecules (usually represented by the osmotic coefficient, g); and (3) the volume of the hydrated protein molecules. We must further assume that there is no change in the state of aggregation with concentration. For protein solutions, this last assumption is usually fulfilled, and the effect of factor (3) is small when we deal with fairly dilute solutions. By measuring the total osmotic pressure, π, over a wide range of protein concentrations, holding the electrolyte concentration (and pH) constant, we may often simultaneously eliminate the effects of the first three factors by extrapolation of π/C values to $C \rightarrow 0$ (where C is the protein concentration in grams of dry protein per liter of solution). This

[8] Donnan, F. G., & Guggenheim, E. A. Zeit. physikal. Chem. 162A: 346. 1932. Adair, G. S. Trans. Faraday Soc. 31: 98. 1935. Donnan, F. G. Trans. Faraday Soc. 31: 80. 1935.
[9] See Cohn, E. J. Chem. Rev. 24: 203. 1939 for references to and discussions of this work.

necessitates making measurements at sufficiently low protein concentrations so as to make the extrapolation to $C \rightarrow O$ as certain as possible, which may be experimentally difficult if the effects of factors (1) and (2) are large. When dealing with proteins at or near their isoelectric point, the ion-difference pressure may be unimportant, but when this is not true, it may be independently evaluated by measurements of membrane potentials, and such measurements are especially valuable since they enable separate estimates of the effects of factors (1) and (2) to be made, and thus provide definite proof as to whether the simple formula

$$M = RT(C/\pi)_{C \rightarrow 0} \tag{2}$$

may be applied (which is possible only when $g \rightarrow 1$ and $p_i \rightarrow 0$ as $C \rightarrow 0$). The applicability of equation (2) can be further studied by making extrapolations of π/C to $C \rightarrow 0$ for various electrolyte concentrations; a value of M thus calculated which is independent of the electrolyte concentration would indicate that the equation could be safely applied.[10]

Measurements of this kind do not give any evidence concerning the mono- or polydisperse nature of the protein under investigation. In case the protein is polydisperse, then one obtains a "number average" molecular weight; that is, $M_n = \Sigma_i (n_i M_i) / \Sigma_i n_i$, but for the substances discussed here this is probably not a serious problem since they are obtainable in monodisperse preparations of high purity. The molecular weights calculated from osmotic pressure by extrapolation to infinite dilution of protein are presumably not seriously affected by even large amounts of hydration. Deviation from a spherical shape also has little effect, except that the deviations from a linear relationship between the osmotic pressure and the concentration usually begin at lower values of the concentration the more asymmetric the molecule. This method suffers the same defect as does the sedimentation-equilibrium method when the time required to reach an equilibrium state is long.

SEDIMENTATION VELOCITY

The measurement of the rate of migration of the protein in a high centrifugal field is one of the most widely used methods for the evaluation of size. Its popularity is mainly based on the fact that such experiments can be performed in a relatively short time, and on the fact that probably this method is the most sensitive to a dis-

---

[10] Analytical proof that $g \rightarrow 1$ when $p_i \rightarrow 0$ and $C \rightarrow 0$ for any arbitrary but constant electrolyte concentration might be possible. See a discussion by Hartley, G. S., & Donnan, F. G. Trans. Faraday Soc. **31**: 106–108. 1935.

tribution of particle size, particularly when used with the newer refractive index methods of analysis of Lamm.[11]

The results obtained by this method are not thermodynamic in nature and are not immediately convertible to values of molecular weight. The result obtained is a "sedimentation constant," s, defined as the rate at which the boundary moves under unit centrifugal field:

$$s = \frac{dx}{dt} \cdot \frac{1}{\omega^2 x} = \frac{d \ln x}{dt} \cdot \frac{1}{\omega^2} \quad (3)$$

This sedimentation constant[12] so obtained can be converted into an estimated size by considering the frictional resistance to sedimentation as evaluated by Stokes' law. For a spherical and unsolvated particle this resistance (on a molar basis) is:

$$f_0 = 6\pi\eta \, Nr = 6\pi\eta N \left[\frac{3\bar{v}M}{4\pi N}\right]^{1/3} \quad (4)$$

or for a non-spherical solvated particle

$$f = 6\pi\eta \, Nr(f/f_0) \quad (5)$$

where (f/f_0) is the frictional ratio which measures the departure of the molecule from the simple behavior, and is thus an indirect method of expressing the asymmetry and solvation of the molecule. Here r is taken as the radius in terms of a spherical molecule of molecular weight M and partial specific volume $\bar{v}$, η the viscosity of the solvent, and N, Avagadro's number. The frictional resistance of a particle can be evaluated from the equation,

$$f = \frac{M(1 - \bar{v}\rho)}{s} \quad (5a)$$

where ρ is the density of the solution. For the sake of convenience, s is usually corrected to the value it would have in pure water at 20° C. by means of the following two equations:

$$s_{20}' = \frac{d \ln x}{dt} \cdot \frac{1}{\omega^2} \frac{\eta_t'}{0.01009} \quad (6a)$$

[11] Lamm, O. Nova Acta Soc. Sci. Upsala IV **10**, (6). 1937; Nature **132**: 820. 1933. See reference 1 p. 253–273.
[12] Values of s obtained by the averaging of individual s values calculated from equation (3) can be shown to be determined only by x values for the initial and the final times, in just the same manner as that described in the note concerning sedimentation equilibrium.[5] Accordingly, we feel that s values obtained by measuring the slope of curves of log x vs. t (η_{20}/η_t), where η_{20}, η_t represents the average viscosity correction during the interval from $t = t_0$ to $t = t$, represent more adequate estimates of the sedimentation constant.

$$s_{20,\,w} = s_{20}' \cdot \frac{\eta_t}{\eta_t'} \cdot \frac{1 - 0.9982\bar{v}_{20}}{1 - \rho_t\bar{v}_t} \tag{6b}$$

where s_{20}' is the value reduced to the protein solution at 20° C., and $s_{20,\,w}$ is the value reduced to water at 20° C. Here η_t', 0.01009 and η_t are the viscosities of the water at $t°$ C., water at 20° C., and the *solvent* at $t°$ C., respectively; and $\bar{v}_{20}$ and $\bar{v}_t$ the partial specific volume of the protein at temperatures of 20° C. and $t°$ C., respectively.

Values of the sedimentation constant s can be used to calculate the molecular weight M if the frictional ratio f/f_0 is known (or assumed), or to calculate f/f_0 if M is known (or assumed):

$$M = \sqrt{162\pi^2\eta^3 N^2 \bar{v}(f/f_0)^3[s/(1 - \bar{v}\rho)]^3} \tag{7}$$

$$f/f_0 = \frac{(1 - \bar{v}\rho)}{6\pi\eta s} \left[\frac{4\pi M^2}{3\bar{v}N^2} \right]^{1/3} \tag{8}$$

In water at 20° C. these equations become:

$$M = 2.45 \times 10^{22}\bar{v}_{20}^{1/2}[(f/f_0)s_{20,\,w}/(1 - 0.9982\bar{v}_{20})]^{3/2} \tag{7a}$$

$$f/f_0 = 1.19 \times 10^{-15}M^{2/3}(1 - 0.9982\bar{v}_{20})/(s_{20,\,w}'\bar{v}_{20}^{1/3}) \tag{8a}$$

The sedimentation constants obtained from measurements on a number of proteins have been found to be a function of the protein concentration, and in order to obtain values suitable for introduction into the above equations, it is necessary to make a suitable extrapolation of the constants observed at finite concentrations.

The accuracy of the extrapolations of observed values of s to values $s_{w,\,20}$ has been tested in various ways. The consistency of values at various temperatures, and in the presence of such foreign materials as urea, alcohol, and heavy water seem to indicate that the viscosity part of Stokes' law is adequate for most systems which, for the moment, we are likely to study[18] This will be further discussed in the next section on diffusion. The effect of the viscosity contribution of the protein is not so well understood, and may be the cause of a large part of the decrease in sedimentation constant with increasing protein concentration. From a theoretical standpoint most workers prefer to use the viscosity of the solvent (including any buffer salts, etc.) rather than of the solution, but certain data would bring this practice into question to some extent. Since the sedimentation constants should in any event be extrapolated to zero protein concentration, it might seem justifiable tentatively to use whichever viscosity leads to the most satisfactory extrapolation function.

---

[18] See, for example, **Svedberg, T.**, & **Eriksson-Quensel, I.-B.** Nature **137:** 400. 1936.

The addition of low concentrations of electrolytes has been shown to be necessary in order to eliminate electrical forces (Donnan effects) in all cases where proteins are transported. Calculations of Tiselius[14] have shown that the addition of relatively small amounts of electrolytes will repress the Donnan effect for most molecules. 0.1 or 0.2 M NaCl or KCl is frequently used at pH values near the isoelectric point, and is adequate for most purposes.

Several other uncertainties in the evaluation of sedimentation constants have been discussed by Svedberg and Pedersen.[1] Perhaps the most important of these is the effect of hydrostatic pressure in the cell upon the density of the solution and the partial specific volume of the protein. The change of viscosity with pressure is small enough to be almost negligible for aqueous solutions under the usual conditions. The boundary disturbances during the accelerating period, as well as those due to diffusion, are of some importance in certain cases, but in general cause no great uncertainty.

DIFFUSION

Measurements of the diffusion constant, D, can be used to evaluate the frictional resistance, f, even more directly than measurements of sedimentation velocity (equation 5), since

$$f = RT/D \tag{9}$$

The molecular weight, M, can be calculated from the diffusion constant if we know f/f_0, or the frictional ratio f/f_0 can be calculated if we know M:

$$M = \frac{R^3T^3}{162\pi^2N^2\eta^3D^3\bar{v}}\,(f_0/f)^3 \tag{10}$$

$$f/f_0 = \left(\frac{RT}{6\pi\eta ND}\right)\left(\frac{4\pi N}{3\bar{v}M}\right)^{1\,3}. \tag{11}$$

In water at 20° C. these equations become:

$$M = 2.42 \times 10^{-14}/[D_{20,\,w}{}^3\bar{v}_{20}\,(f/f_0)^3] \tag{10a}$$

$$f/f_0 = 2.89 \times 10^{-5}/[D_{20,\,w}(\bar{v}_{20}\,M)^{1/3}] \tag{11a}$$

where $D_{20,\,w}$ is the diffusion constant (cm²/sec.) corrected to water at 20° C. by means of the equation:

$$D_{20,\,w} = D \cdot \frac{293.2}{T} \cdot \frac{\eta_t}{\eta_{20,\,w}} \tag{12}$$

[14] Tiselius, A. Kolloid. Zeit **59**: 306. 1932. See reference 1, p. 23–8

The measurement of the diffusion constant can be carried out either by "free diffusion" (Lamm & Polson)[15] or by use of sintered glass discs (Northrop & Anson).[16] Both methods are fairly widely used. The former method is an absolute measurement involving no primary standards, and is the most advantageous in most cases. The sintered glass discs, on the other hand, require a calibration of some sort to evaluate the pore size, and the selection of a primary standard is not an easy task. Also there is little evidence obtainable from the latter method in regard to the mono-disperse nature of the material and to the symmetry of the diffusion boundary, both of which are of considerable importance in the study of an unknown molecule.

In the free diffusion method we usually evaluate the quantity $\dfrac{dc}{dx}$ as a function of the distance x by means of a scale photograph or Schlieren method, and we may then evaluate the diffusion constant in several ways.[15, 17] The two most commonly used, and probably the most precise are (1) the measurement of the width of the peak at a point where $\dfrac{dc}{dx}$ (or z) is equal to $\dfrac{1}{e} \cdot \dfrac{dc}{dx}$ (or $z_{max}./e$), and (2) the measurement of the maximum height of the peak and the area enclosed by the peak. The agreement between these methods is usually of the order of 3–5%. When dealing with monodisperse systems it is usually found that the curves representing the distribution of protein as a function of the distance agree well with ideal distribution curves, and this method has been used as an indication of the monodisperse nature of certain protein preparations. Deviations can likewise be used to evaluate the degree of polydispersity of a material. In certain cases it is observed that the diffusion curve is asymmetrical. The reasons for this behavior are not altogether clear, and cannot be explained simply by the presence of several components. The calculations of diffusion constants under these conditions are necessarily very uncertain.

The adequacy of the viscosity and temperature corrections can best be demonstrated by several measurements of diffusion constants at temperatures between 25° C. and 5° C.[18] When these two sets of data are corrected at 20° C., they are in excellent agreement. The viscosity effect has also been studied by variation of the salts and other molecules of the solvent.

[15] Lamm, O., & Polson, A. G. Biochem. Jour. 30: 528. 1936. Lamm, O. Nova Acta Soc Sci. Upsala IV 10: (6). 1937.

[16] Northrop, J. H., & Anson, M. L. Jour. Gen. Physiol. 12: 543. 1929.

[17] Williams, J. W., & Cady, L. C. Chem. Rev. 14: 171. 1936.

[18] Polson, A. G. Kolloid. Zeit. 87: 149. 1939. Mehl, J. W. unpublished work.

SEDIMENTATION VELOCITY AND DIFFUSION

The equations (7) and (8) for sedimentation velocity, and (10) and (11) for diffusion, involve both the molecular weight M and the frictional ratio f/f_0. By their proper combination, however, we obtain two new equations which uniquely determine these quantities:

$$M = RTs/[D(1 - \bar{v}\rho)] \tag{13}$$

$$f/f_0 = \frac{1}{6\pi\eta}\left(\frac{RT}{DN}\right)^{2/3}\left(\frac{4\pi(1 - \bar{v}\rho)}{3\bar{v}s}\right)^{1/3} \tag{14}$$

In water at 20° C. these equations become:

$$M = 2.44 \times 10^{10}s_{20, w}[D_{20, w}(1 - 0.9982\bar{v}_{20})] \tag{13a}$$

$$f/f_0 = 1.00 \times 10^{-8}[(1 - 0.9982\bar{v}_{20})/D_{20, w}{}^2 s_{20, w}\bar{v}_{20})]^{1/3} \tag{14a}$$

Although equation (13) is not thermodynamic in nature, it can be shown that at least in most cases the values of M so obtained are not affected to any large extent by hydration, and represent estimates just as reliable as those obtained from sedimentation equilibrium.[19, 20] Values so obtained are independent of shape. The main difficulty encountered is in the selection of values of s and D for infinite protein dilution, which should be used in these equations. Values at finite concentrations do not in general vary enormously from these infinite dilution values, but small errors can be introduced. If s and D vary with concentration in the same manner, so that the ratio s/D is constant, values of s and D taken under similar conditions may be used in equation (13). Values at infinite dilution must be obtained, however, before f/f_0 values can be obtained.

Values of the frictional ratio f/f_0 greater than unity may be explained on the basis of either the asymmetry or the hydration of the protein molecules, or, more probably, a combination of these two effects. Thus the observed frictional ratio f/f_0 can be broken into two factors:

$$f/f_0 = (f/f_e)(f_e/f_0)$$

where the first denotes the influence of hydration, and the second the influence of asymmetry. As a first approximation, necessary for the mathematical treatment of the problem, we may assume that protein

[19] Lamm, O. ark. Mat. Astrom. och. Fysik. 21: 1. 1929. Lansing, W. D., & Kraemer, E. O. Jour. Am. Chem. Soc. 58: 1471. 1936. Reference 1, pp. 20–23, 62–66.

[20] It would appear that sedimentation velocity-diffusion estimates of molecular weights are even more reliable than the sedimentation equilibrium values for most of the molecules discussed in this conference, because of the somewhat smaller probable errors in the determination of diffusion and sedimentation constants in comparison with probable errors of the sedimentation equilibrium method.

molecules can be treated by an ellipsoid of revolution. Using the equations derived by Perrin[21] for the frictional ratio of ellipsoids of revolution, both elongated (prolate) and flattened (oblate) we may calculate values of f_e/f_0 to be expected from various axial ratios, a/b of the ellipsoid. Here a is taken as the half-length of the ellipsoid along the geometric axis of revolution, and b is the equatorial radius.
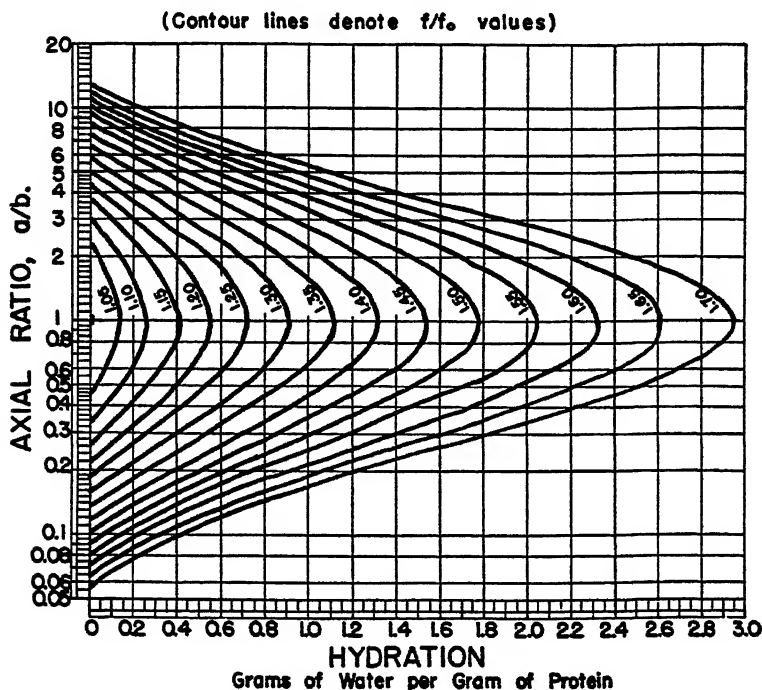

FIGURE 1. Values of axial ratio and hydration in accord with various frictional ratios (contour lines denote f/f_0 values).

The contribution due to hydration, f/f_e, can be calculated from the equation (see reference 1, p. 65)

$$f/f_e = (1 + w/\bar{v}\varrho)^{1/3}$$

where w is the number of grams of water bound by one gram of protein. A figure can then be drawn, using hydration as ordinate and axial ratio as abscissa (on a logarithmic scale), in which various values of the frictional ratio are given by various contour lines. Such a diagram

[21] Perrin, F. Jour. Phys. Radium (7) 7: 1. 1936. Hertzog, R. O., Illig, R., & Kuder, H. (Zeit. physikal. Chem. A167: 329. 1933) have obtained results which are identical with those of Perrin. See reference 1 pp. 40–42.

is given in FIGURE 1.[22]   This can be used for the calculation of all amounts of hydration and asymmetry which might combine to yield a given frictional ratio.

VISCOSITY

Another method which is capable of giving data in regard to the shape and hydration of protein molecules is the study of viscosity. The basis of this method is somewhat less secure from both a theoretical and experimental point of view than the other methods already discussed.   The original derivation by Einstein of an equation relating the viscosity to the volume fraction, Φ, of a suspension of rigid spheres whose size is large compared to the size of the solvent molecules,

$$\eta/\eta_0 - 1 = F\Phi$$

where F is equal to 2.50, has been modified by a number of more recent investigations, and can be applied to ellipsoidal shapes.   Equations by Onsager,[23] Guth,[24] Burgess,[25] Peterlin,[26] and Simha[27] for the case of ellipsoids, and of Kuhn[28] and Huggins[29] for long rods, are at present available to us, and modify the equation of Einstein in that they give values for the parameter F which are functions of the axial ratio.   For molecules of low asymmetry, as is the case of nearly all the molecules under discussion here, the treatments of ellipsoids are more satisfactory, and the solution of Simha seems to be the most adequate.

The volume fraction, Φ, can be expressed in terms of the protein concentration, c, (grams dry protein per liter of solution), the partial specific volume of the anhydrous proteins molecules, $\bar{v}$, the number of grams of solvent bound by one gram of dry protein, w, and the density of the solvent, ρ.   We may thus rewrite the above equation

$$\eta/\eta_0 - 1 = F\bar{v}c\,[1 + w/(\bar{v}\rho)]/1,000.$$

Using this equation and values of F obtained from Simha's equation,[30] we may now draw another diagram similar to FIGURE 1, but with con-

[22] We have set $1/(\bar{v}\rho)$ equal to 1.34 for this calculation, a value which can be used for most proteins.

[23] Onsager, L.   Phys. Rev. **40**: 1029.   1932.

[24] Guth, E.   Kolloid. Zeit. **74**: 147.   1936.

[25] Burgess, J. M.   "Second Report on Viscosity and Plasticity."   Amsterdam.   1938. Chapter 3 (For cylinders).

[26] Peterlin, A.   Zeit. physik. **3**: 232.   1938; Kolloid. Zeit. **86**: 230. 1939   (Ellipsoids).

[27] Simha, R.   Jour. Phys. Chem. **44**: 25.   1940.

[28] Kuhn, W.   Zeit. physikal. Chem. **A161**: 1.   1932.

[29] Huggins, M. L.   Jour. Phys. Chem. **42**: 911.   1938; **43**: 439.   1939.

[30] Mehl, J. W., Oncley, J. L., & Simha, R.   Science **92**: 132. 1940 give tables of F (ν in their notation) as a function of the axial ratio $a'b$ ($1/\rho$ in their notation) for both elongated (oblate) and flattened (prolate) ellipsoids of revolution.
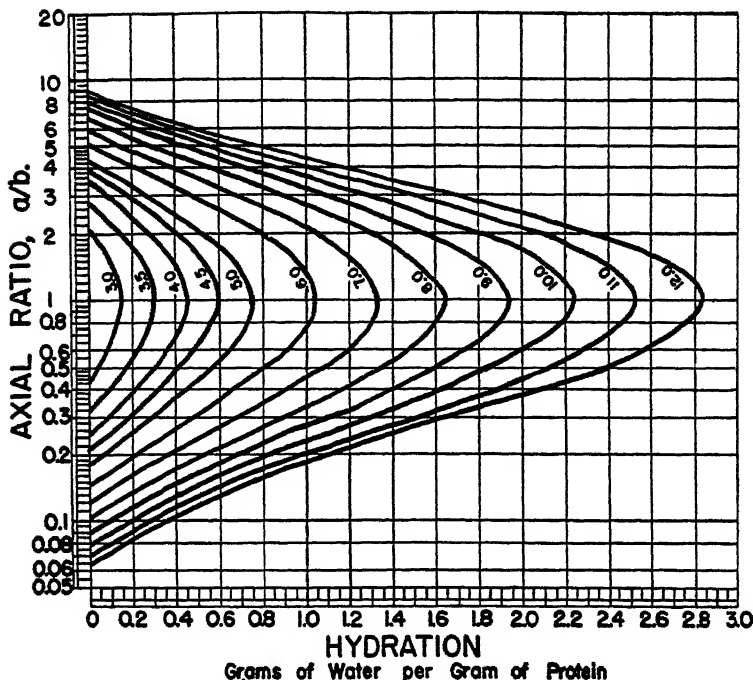
FIGURE 2. Values of axial ratio and hydration in accord with various viscosity coefficients (contour lines denote $(\eta/\eta_0 - 1)\, 1000/\bar{v}c$ values).

tour lines for various values of $(\eta/\eta_0 - 1)\, 1{,}000/\bar{v}c$. Such a diagram is shown in FIGURE 2.

In general it is found that the specific viscosity, $\eta/\eta_0 - 1$, is a linear function of the concentration only at very low concentrations (less than 1% for most of the molecules under discussion). The available data is best treated by an extrapolation of $(\eta/\eta_0 - 1)/c$ vs. concentration curves to infinite dilution.[31] The values so obtained in various investigations are not in very good agreement, and show the need for

[31] Polson, A. G. Nature 137: 740. 1936; Kolloid Zeit. 88: 51. 1939 has used another method for finding the limiting value of $(\eta/\eta_0 - 1)/c$. If we used the empirical equation of Arrhenius, $\eta/\eta_0 = \alpha^c$, we see that[3] $d(\eta/\eta_0)/dc = \alpha c \ln \varepsilon$, and the limit of this quantity at infinite dilution ($c = 0$) is $\ln \alpha$. Thus it is only necessary to evaluate $\ln \alpha$ from the observed data, and extrapolate this quantity to low concentration if any concentration dependence is observed. This quantity $\ln \alpha$ can easily be calculated from the equation:

$$1/c \ln \eta/\eta_0 = \ln \alpha$$

This method should be reliable if the results so obtained are extrapolated to zero concentrations. If used for finite concentrations, it involves the assumption that the Arrhenius equation is correct. When extrapolated to zero concentrations, the values obtained by this method agree with those obtained by the extrapolation of the quantity $(\eta/\eta_0 - 1)/c$.

more precise measurements of the viscosities of very dilute protein solutions.

The problem of obtaining more accurate viscosity data in dilute protein solutions is a serious one. The effect of surface denaturation in solutions of dilute proteins is well known and may be responsible for some of the lack of agreement between workers. The effect of pH on electrolyte-free protein solution is of great importance, and probably the best procedure for viscosity measurements is to conduct them in the presence of about 0.2N salt. Polson has shown[32] that even large pH changes fail to introduce errors in viscosity under these conditions.

The effect of the velocity gradient used in determining the viscosity is another possible source of error, particularly when dealing with highly asymmetric molecules. It can easily be shown that the gradients usually used in capillary viscometers, in the study of the proteins being discussed here are much too small to be of importance here, however.[33]

DIELECTRIC DISPERSION

The study of dispersion of the dielectric constant is capable of giving us data in regard to the size and shape of molecules in addition to information concerning the dipole moment. This problem was discussed during the recent dielectric symposium of the Academy,[34] particularly from the point of view of the dipole moments of protein molecules. For the present discussion, it is probably best to think of these data in terms of a rotary diffusion constant, Θ, analogous to the linear diffusion constant, D. The experimental data required are dielectric constants measured for suitable solutions of low conductivity over a fairly wide frequency range. By applying the equations of Debye, as modified by recent treatments of Wyman, Onsager, Perrin and others, we can analyze measurements of dispersion of the dielectric constant into various critical frequencies, ν_c, relaxation times, $\tau(\tau = 1/2\pi\nu_c)$ or rotary diffusion constants, $\Theta(\Theta = \pi\nu_c)$. Having these critical frequencies, we may apply Perrin's equations and calculate values of a hypothetical critical frequency, ν_0, for a spherical molecule of equal volume, from which we can evaluate the size of the molecule, and the axial ratio a/b, of an assumed ellipsoidal shape which may be either a flattened or an elongated ellipsoid of revolution.

[32] Polson, A. G.   Kolloid. Zeit. 88: 51.   1939.

[33] For a discussion of this error see Kroepelin, J.   Kolloid. Zeit. 47: 300.   1929.   It has been briefly discussed by Edsall, J. T., & Mehl, J. W.   Jour. Biol. Chem. 133: 418.   1940. See also Polson.[32]

[34] Oncley, J. L., Ferry, J. D., & Shack, J.   Ann. N. Y. Acad. Sci. 40: 371.   1940.

The method briefly reviewed above has been applied to a considerable number of the proteins under discussion here. In the case of egg albumin, serum albumin, edestin and possible lactoglobulin, it was found that two critical frequencies were required to reproduce the data obtained. In such cases, a comparison of experimental results with calculations from Perrin's equation give us values for v_0, ⊛ and a/b if we assume that the molecule can be represented by an ellipsoid of revolution. Here ⊛ is the "dipole angle" of the molecule, defined as the angle between the a axis of the assumed ellipsoid and the dipole moment vector. The resulting size of the molecule is calculated from v_0 (or τ_0) and gives us a molecular volume, V, for the protein as it exists in the solution (including whatever water of hydration—w grams per gram of dry protein—may be considered rigidly bound to the molecule. This molecular volume can then be compared with $M\bar{v}$ and the amount of hydration computed:

$$V = RT/6\pi\eta v_0 = RT\tau_0/3\eta = M\bar{v} \, (1 + w/\bar{v}\rho)$$

In the case of hemoglobin, lactoglobulin, and insulin, we find that the data can be well represented by a single relaxation time. This can be explained either by the assumption 1) of a spherical molecule (in which case $a/b = 1$ and $v_c = v_0$); 2) that the other critical frequency is associated with a dipole moment of such a small magnitude that it is not observed, or 3) that the ellipsoid is of flattened shape. In any of these circumstances the observed critical frequency, v_c, may be either v_a, the critical frequency associated with the a axis, or v_b, associated with the b axis. It is then necessary to calculate all possible values of v_0, ⊛ and a/b, and then evaluate all possible combinations of axial ratio a/b and hydration w which could be used to explain the observed dielectric dispersion curves.

The simple picture presented in the above paragraph is unfortunately only a first approximation. Although two critical frequencies can be chosen to represent the large portion of the dielectric constant versus log frequency curves, the agreement between calculated and observed curves is not as good as could be desired. This deviation is of a form to cause more of a smoothing in the case of the experimental curves than would be calculated. The choice of even a third critical frequency does not improve the agreement to any appreciable extent. A more likely explanation of this effect, however, can be suggested.

In a solution of protein at its isoelectric point we would expect a large number of dipole species of identical size and shape, but of different electrical symmetry. A number of such species will depend

upon the number of dissociating groups and their dissociation constants.[35]　If we express the electrical symmetry in terms of electric moments μ_a and μ_b along the two axes of an ellipsoid of revolution we see that each species may have different values of these quantities, depending upon the distribution of charged groups on the protein molecule.　For a spherical molecule this leads to no additional complications, and an average moment is obtained.　For an asymmetric molecule, on the other hand. each species should make a contribution to each of the separate dispersions. and the observed curve would be an average of many curves of varying shapes, even though ν_a and ν_b remained constant, and would thus have lost much of the sharpness of the individual curves.　Zein seems to be a case in which the shape can be rather simply calculated, since it contains so few dissociating groups that at the isoelectric point there is probably a fairly simple behavior.[36]　Edestin, however, is perhaps a case where the data must be treated in the more complex manner discussed here.　Horse serum pseudoglobulin-γ may be another case of this type.

The measurements of dielectric constant at frequencies above the dispersion region can also be used to calculate the hydration of proteins.[37]　The dielectric constant obtained under these conditions is due almost entirely to solvent dipoles, and if we take the difference, $\Delta\epsilon_\infty$, between the dielectric constant so measured, ϵ_∞, and that of the pure solvent, ϵ^0, we may use it to evaluate the amount of water associated with each gram of anhydrous protein.　Only a few measurements of this type are available.

DOUBLE REFRACTION OF FLOW

The measurement of the degree of orientation of molecules as measured by the double refraction of the solution caused by velocity gradients in the fluid has been attempted for many systems, and the data thus obtained can be used in the evaluation of the molecular asymmetry, by means of the rotary diffusion constant, Θ.[38]　Such methods have been proven to be of great value in the study of some of the larger protein molecules. such as myosin and tobacco mosaic virus,

[35] Edsall, J. T., & Blanchard, M. H.　Jour. Am. Chem. Soc. **55**: 2337.　1933.　Greenstein, J. P.　Jour. Biol. Chem. **93**: 479.　1931.　Linderstrøm-Lang, K.　Compt. rend. trav. lab. Carlsberg **15** (7): 1924.　See also Shack, J.　Ph.D. Dissertation, "Dielectric Absorption of Protein Solutions."　Harvard University 1939.

[36] Elliott, M. A., & Williams, J. W.　Jour. Am. Chem. Soc. **61**: 718.　1939.　See also Wyman, J.　Jour. Biol. Chem. **90**: 443.　1931; and Watson, C. C., Arrhenius, S., & Williams, J. W.　Nature **137**: 322.　1936.

[37] Oncley, J. L.　Jour. Am. Chem. Soc. **60**: 1122. 1938, note 43.

and have been adequately discussed elsewhere.[39]   Recently there are available measurements on one of the molecules under discussion here: serum albumin solutions of various solubility have been investigated in glycerol-water mixtures by Sadron,[40] and he concludes that serum albumin fractions of low solubility are fairly symmetrical.   The case of the more soluble albumins was complicated by the presence of molecules of varying amounts (and sign) of birefringence.   Measurements on egg albumin by Boehm and Signer[41] showed no large degree of asymmetry.   The application of this method to other proteins is being undertaken in our laboratory by Edsall and Mehl, and it is hoped that asymmetries can be obtained from this method which may be compared with those deduced on the same preparations by the other methods that have been discussed.

PARTIAL SPECIFIC VOLUME

The partial specific volume of the protein is involved in nearly all of the experiments which have been considered, has been used directly by Adair and others to measure the hydration of proteins, and is used in certain of the X-ray calculations of molecular weight discussed in another review of this conference.[42]   This quantity is evaluated from measurements of density of solutions of the protein under investigation.[43]   It has been pointed out in the introduction to this conference that its value varies somewhat for various proteins, but is directly related to the amino acid composition of the protein.[44]   Many proteins have values between 0.745 and 0.750, but there are many cases falling outside this range, and the occasional practice of using specific volumes of related proteins, rather than carrying out the necessary density measurements, should be discontinued.   This is especially true when it occurs in the function $(1 - \bar{v}\rho)$, since errors in $\bar{v}$ are magnified approximately fourfold.

It has been pointed out that molecular weights as estimated by sedimentation-equilibrium or osmotic pressure are thermodynamic methods which are unaffected by any hydration of the protein.   Like-

[38] The theoretical side of this problem has been treated by Boeder, P., Zeit. Phys **75**: 258.   1932.   **Kuhn, W.**   Kolloid. Zeit. **62**: 269.   1933.

[39] von **Muralt, A. L.**, & Edsall, J. T.   Jour. Biol. Chem. **89**: 315   1930.   **Mehl, J. W.** Cold Spring Harbor Symp. Quant. Biol. **6**: 218.   1938.   **Lauffer, M. A.**, & **Stanley, W. M.** Jour. Biol. Chem. **123**: 507.   1938; Jour. Phys. Chem. **42**: 935.   1938.

[40] **Sadron, C.**, Bonot, A., & **Mosimann, H.**   Jour. chim. phys. **36**: 78.   1939.

[41] **Boehm, G.**, & **Signer, R.**   Helv. chim. acta **14**: 1370.   1931.

[42] **Fankuchen, I.**   Ann. N. Y. Acad. Sci., this number.

[43] See **Kraemer**'s section of Reference 1, pp. 57-62, for details of this measurement.

[44] **Cohn, E. J.**   Ann. Rev. Biochem. **4**: 93. 1935; Ann. N. Y. Acad. Sci., this number.

wise the sedimentation-diffusion method gives molecular weights certainly very little affected by even very large amounts of hydration. This can best be explained by the fact that the factors $M_a(1 - \bar{v}_a \rho)$ and $M_w(1 - \bar{v}_{wc} \rho)$ are equal, where the subscripts a refer to properties of the anhydrous molecules and the subscripts w to those of the hydrated molecules. This problem has been discussed at length by Lansing and Kraemer,[45] who also derive equations to show that these same conditions hold at least approximately for three-component systems with solvation.

Adair and others[46] have attempted to study the amount of hydration by the measurement of protein crystal densities. They calculate the number of grams of water associated with one gram of dry protein, w, from the formula

$$w = (\bar{v}_c - \bar{v}_a)/(1/\rho - \bar{v}_c).$$

where $\bar{v}_c$ is the specific volume of the protein crystals, $\bar{v}_a$ is the specific volume of anhydrous protein, and ρ is the density of water. They measured $\bar{v}_c$ by immersion in solutions of suitable density to just suspend the protein crystals. These solutions consisted of such materials as sucrose, ammonium sulfate, mono-sodium phosphate and potassium citrate in aqueous solution. The objections to these methods are (1) that they assume that the crystals consist of two components; A, composed of anhydrous protein and water, and B, composed of dispersion medium; (2) that the environment of the crystals is so changed in the process of obtaining a suspending media of proper density that the densities so obtained are no longer those of the protein crystals under normal conditions, and (3) that the "water of hydration" may not have the same density as water in its usual state. (This third objection is equally true of all the methods described for the calculation of hydration from values of molecular volumes.)

MOLECULAR WEIGHT

It remains to attempt an evaluation of the data available for the proteins chosen for discussion in this conference. The measurements to be considered are compiled in TABLE 1, and the accompanying notes. X-ray estimates of molecular weight are considered by Fankuchen.[42] The case of each of the molecules can perhaps be considered individually.

[45] Lansing, W. D., & Kraemer, E. O. Jour. Am. Chem. Soc. **58**: 1471. 1936. See also reference 1, pp. 62–66.

[46] Adair, G. S., & Adair, M. E. Proc. Roy. Soc. B120: 422. 1936. See also Adair, G. S. Proc. Roy. Soc. B127: 18. 1939.

Insulin

All the available ultracentrifuge data on insulin were obtained prior to the recent revision of sedimentation equilibrium values. The value of 35,100 by sedimentation equilibrium may therefore be changed. The diffusion constant has been measured by the latest methods, and the sedimentation velocity constants have not appreciably changed since 1931. The value of 41,000 may thus tentatively be taken as the more likely of the two given here.

Lactoglobulin

The available data on lactoglobulin are all of quite recent origin. A value of 40,000 ± 2,000 is in good agreement with both of these results.

Chymotrypsin and chymotrypsinogen

These proteins have been crystallized by Northrop and his co-workers, who have estimated molecular weight by osmotic pressure methods. The ultracentrifugal molecular weights of these substances are being studied by J. W. Williams and his co-workers, and additional data will soon be available, for their more accurate evaluation.

Pepsin

The values for pepsin are of recent origin, unpublished to a large extent, and all fall within a value of 37,000 ± 2,000.

Egg Albumin

Choice of a value for the molecular weight of egg albumin is at present somewhat difficult. Both the sedimentation velocity and diffusion constants seem quite well established, and lead to a molecular weight close to 44,000, somewhat higher than that given by the equilibrium ultracentrifuge. Sörensen in 1917 estimated the molecular weight of this protein to be 34,000 from osmotic pressure measurements. Adair[47] in speaking of the osmotic pressure results refers to "a recalculation based on Sörensen's data made by Adair which gave the value 43,000. The same figure was obtained by Marrack and Hewitt from measurements of osmotic pressure. Taylor, Adair and Adair using the same method obtained the value 46,000 for material which had been recrystallized six times."

[47] Adair, G. S. Ann. Rev. Biochem. 6: 180. 1937.

Serum Albumin (Horse)

All of the recent data on this protein seem in good agreement with the figure 70,000 ± 3,000. The new fractionation of serum albumin by McMeekin[48] reopens this problem to a certain extent.

Hemoglobin (Horse)

We may take 66,700, the value as obtained from chemical analysis, for the molecular weight of hemoglobin. The value from sedimentation and diffusion is about 5% lower than this figure, but is complicated by the fact that extrapolations are difficult to make because dissociation seems to occur at concentrations below 0.8 per cent, and that the optical aanlysis of hemoglobin solutions by the refractive index methods are somewhat less precise than is usually the case because the photographic work must be done in the red or yellow portion of the spectrum.

Edestin

Until further details are published, we probably should accept the value of 310,000 for the molecular weight of this protein. A new determination by sedimentation equilibrium would be of considerable value in confirming the newer result from sedimentation velocity and diffusion.

Excelsin

In view of the difficulties involved in the osmotic pressure measurements in the case of such a large molecule the value from sedimentation and diffusion must be considered the most accurate that we possess, though it is to be hoped that it, like edestin, will again be studied in the equilibrium ultracentrifuge.

Bushy Stunt Virus

The molecular weight of this material is best taken as 7,600,000, on the basis of sedimentation equilibrium measurements. It would be interesting to have a direct measurement of the diffusion constant of this virus.

ASYMMETRY AND HYDRATION

The calculation of asymmetry by most of the methods that we have discussed is complicated by the effect of hydration. In order to include this effect, we have attempted to present the data on asymmetry and hydration in graphic form, the same coordinates being adopted as

---

[48] McMeekin, T. L. Jour. Am. Chem. Soc. **61**: 2884. 1939; **62**: 3393. 1940.
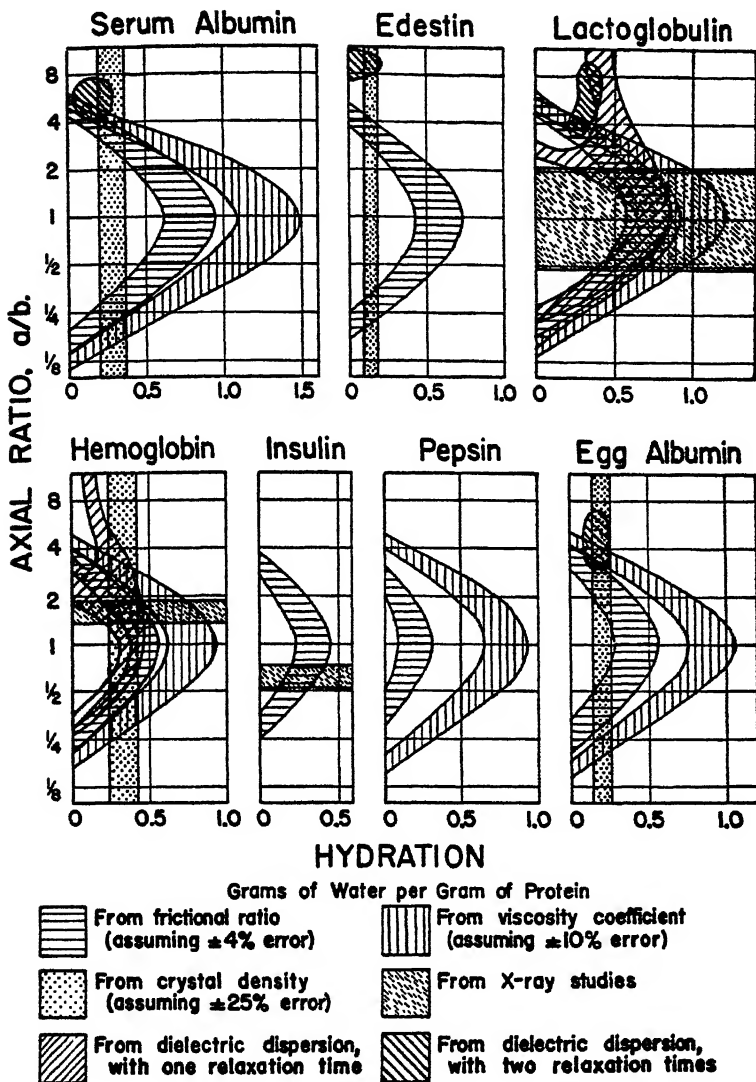
FIGURE 3. Asymmetry and hydration of certain protein molecules.

were previously employed in FIGURES 1 and 2, to present the data from diffusion, viscosity, dielectric dispersion, crystal density and X-ray diffraction studies (FIGURE 3). The diffusion and viscosity results were plotted by taking the values for f/f_0 and $(\eta/\eta_0 - 1)1000/\bar{v}C$ recorded

in TABLE 1, and arbitrarily choosing probable experimental errors of ± 4 per cent for f/f_0 and ± 10 per cent for $(\eta/\eta_0 - 1) - 1000/\bar{v}C$. Results from dielectric measurements were calculated from the relaxation times recorded in TABLE 1, with as reasonable estimates of probable error as possible. The results from crystal density measurements are arbitrarily assumed to be $\pm 25\%$, and are taken from Adair and Adair.[46] The sources of X-ray data, and their reliability, are discussed by Fankuchen.[42, 49] By a study of the diagrams of FIGURE 3, we can best see just how well the various methods seem to agree.

Insulin

Studies of diffusion-sedimentation velocity and X-ray diffraction have been made.[50] The X-ray results seem to strongly favor a flattened ellipsoidal model, and the values of $a/b = 0.6$, $w = 0.2$, seem in good agreement with both methods.

Lactoglobulin

Results from diffusion-sedimentation velocity, viscosity, dielectric dispersion, and X-ray diffraction, are available. The method of dielectric dispersion seems to show a slight deviation from a single relaxation time curve, and the values $a/b = 4$, and $w = 0.3$ would appear most likely from this work; these values would agree with diffusion and viscosity data but are in poor agreement with X-ray results. If the small deviation suggested by the dielectric dispersion is neglected, then values of a/b ca. 1, and $w = 0.6$, taken from the X-ray results, would be in agreement with the single observed relaxation time, as well as all other data, but there is no data to support such large estimates of hydration.

[49] We wish to acknowledge the cooperation of Dr I. Fankuchen in the preparation of the X-ray estimates in FIGURE 3, and the privilege of seeing the manuscript of several unpublished papers of Drs. Dorothy Crowfoot and Dennis P. Riley, and the Ph.D. Dissertation (University of Cambridge, 1939) of Dr. M. F. Perutz, entitled "The Crystal Structure of Horse Methaemoglobin." The values for asymmetry as calculated from X-ray measurements depend upon 1) the distribution of the molecular centers; 2) the assumption that the molecules closely approximate ellipsoids of revolution; and 3) the arrangement chosen for the packing of these ellipsoids. The estimation of the distribution of the molecular centers (1) depends upon several factors: for crystals containing a small number of molecules per unit cell (insulin and hemoglobin) this distribution of molecular centers is fixed by the symmetry of the crystal, whereas for crystals containing large numbers of molecules per unit cell (such as lactoglobulin) this distribution can be determined only by a complicated study of the X-ray intensities. The choice of arrangement (3) is influenced by such guides as symmetry considerations, the character of peaks in Fourier projections, etc., and may be subject to considerable error.

[50] Dielectric dispersion results in propylene glycol, the only solute in which insulin has thus far been studied by this method, cannot be explained on the basis of a molecular weight of 40,000 (See appendix to Table I).

Pepsin

The data from diffusion-sedimentation velocity and from viscosity are in poor agreement, and the interpretation of X-ray results is uncertain and has not been considered. No most probable values can be suggested from the available data.

Egg Albumin

Diffusion-sedimentation, viscosity, dielectric dispersion, and crystal density studies are available for egg albumin. The values $a/b = 4$ and $w = 0.2$ would seem to agree quite well with all of the available data.

Serum Albumin (Horse)

Results from diffusion-sedimentation, viscosity, dielectric dispersion, and crystal density have been considered. The values $a/b = 5$, and $w = 0.2$ agree well with these data, but it must be remembered that these various measurements have been made on serum albumin fractions that were prepared by several methods and may have had varying axial ratios and hydrations.

Hemoglobin (Horse)

This protein has been studied by all of the available methods, and values of a/b ca. 1.6, $w = 0.3$ seem in good agreement with these data. There is some uncertainty in the f/f_0 value, however, as discussed in the appendix of TABLE 1.

Edestin

Measurements of crystal density, dielectric dispersion and diffusion-sedimentation velocity are available, and are not in very good agreement. The f/f_0 value chosen may be too low, and the value of a/b from dielectric dispersion might be too large. X-ray studies are difficult to interpret and have been omitted from consideration. An elongated molecule with w about 0.1 and a/b between 4 and 8 would appear most probable.

Chymotrypsin, Chymotrypsinogen, Excelsin, and Bushy-Stunt Virus

These molecules have been studied only by diffusion (f/f_0) and thus cannot be critically examined.[51] The axial ratio vs. hydration diagrams can be obtained from FIGURE 1, and are not redrawn.

---

[51] Some X-ray data do exist on chymotrypsin, but their interpretation in terms of asymmetry is uncertain.

NOTES FOR TABLE 1

Sedimentation Equilibrium

INSULIN.—The average of three values from Sjögren and Svedberg (49). Svedberg and Pedersen state (57) that these results give ". . . probably too low a value, since not sufficient time was allowed for the equilibrium to be established."

LACTOGLOBULIN.—The average of about sixteen experiments with extremes from 36,000 to 42,500 from Pedersen (38).

PEPSIN.—From Eriksson-Quensel (13). Philpot and Eriksson-Quensel (41) gave 33,500 as the mean of two experiments.

EGG ALBUMIN.—From Pedersen (39). Svedberg and Pedersen (57) state, "When the light-adsorption method was used, values from 40,000 to 46,000 were generally found, but the agreement between the different experiments was not very good. With the refractive index method the single experiments agreed better, and as a mean value, $M_e = 40,500$ was found." This value replaces a former value of about 34,500 obtained from work of Svedberg and Nichols (55), Nichols (32), Sjögren and Svedberg (48) and Svedberg and Eriksson (53), since, "It seems . . . as if the duration of their experiments have been too short to secure real sedimentation equilibrium, so that the molecular weight calculated by them is too small" (57).

SERUM ALBUMIN.—From Pedersen (39). This value checks well with an older value of $68,150 \pm 2,000$ obtained by Svedberg and Sjögren (58).

HEMOGLOBIN.—From Pedersen (39).

TABLE 1

SIZE AND SHAPE OF CERTAIN PROTEIN MOLECULES*

	Insulin	Lacto-globulin	Chymo-trypsin	Chymo-trypsin-ogen	Pepsin	Egg Albumin	Serum Albumin	Hemo-globin	Edestin	Excel-sin	Bushy-Stunt Virus
Molecular Weight											
Equilibrium Centrifuge	35,000	38,000	41,000	36,000	39,000	40,500	68,000	68,000	**	*·	7,600,000
Osmotic Pressure					35,000	34,000–46,000	73,000	67,000		214,000	
Sedimentation-Diffusion	41,000	41,500		38,000	35,500	44,000	70,000	63,000	310,000	295,000	
Sedimentation Constant* $s^0_{20}, w \times 10^{13}$	3.5	3.12		3.1	3.3	3.55	4.46	4.41	12.8	13.3	146.
Diffusion Constant $D^0_{20}, w \times 10^7$	8.2	7.3	7.1	7.9	9.0	7.8	6.1	6.9	3.93	4.26	
Partial Specific Volume, $\bar{v}$	0.749	0.751			(0.750)	0.749	0.748	0.749	0.744	0.743	0.739
Relaxation Times, $\tau^0_{20}, w \times 10^8$	1.7	18; 5.6				18; 4.7	36; 7.5	8.4	250; 28		
Viscosity Coefficient $(\eta/\eta_0 - 1) 1000/(\bar{v}C)$		6.0			5.2	5.5	6.8	5.1			
Frictional Ratio, f/f_0	1.13	1.26	1.30	1.20	1.08	1.16	1.27	1.16	1.21	1.13	1.09

* X-ray measurements have been made on tobacco seed globulin in addition to the other molecules recorded here. The only measurement on this molecule which should be recorded here was $s \times 10^{13} = 12.7$. (See notes to TABLE 1.)

** See accompanying notes.

EDESTIN.—A value of 208,000 was given by Svedberg and Stamm (60), but is no longer given in tables published from Upsala. Svedberg and Pedersen (57) state that ". . . the duration of these runs was too short to secure equilibrium. . ."

EXCELSIN.—A value of 212,000 was given by Svedberg and Sjögren (59), but is subject to the same remarks just made concerning edestin.

BUSHY-STUNT VIRUS.—From McFarlane and Kekwick (26). It represents an average from three experiments which vary between extremes of 7,100,000 and 8,100,000.

Osmotic Pressure

CHYMOTRYPSIN.—From Kunitz and Northrop (22).

CHYMOTRYPSINOGEN.—From Kunitz and Northrop (22).

PEPSIN.—From Northrop (33).

EGG ALBUMIN.—Sörensen (50) gave 34,000 in aqueous solution, a value which has been revised by Adair (3) to give 43,000. Burk and Greenberg (9) gave 36,000 in aqueous solution and 36,000 in urea solution. Marrack and Hewitt (24) gave 43,000 in aqueous solution, and Taylor, Adair and Adair (61) gave 46,000.

SERUM ALBUMIN.—Sörensen (51) gave 45,000 which Burk (7) showed was due to the deviation from an ideal solution. When corrected for this deviation, Burk gave 76,000 from these measurements, Adair (1) gave 62,000, Adair and Robinson (4) gave 72,000. Burk (7) found 75,000 and a similar value (73,000) for denatured material. Roche, Dourier and Marquet (46) gave 69,000 for human serum albumin.

HEMOGLOBIN.—Adair (2) gave 67,000 in aqueous solution. Burk and Greenberg (9) gave 66,500 in 6.5 M glycerol solution. Values obtained in urea are approximately one-half of these values.

EXCELSIN.—From Burk (8).

Sedimentation Velocity and Diffusion

INSULIN.—$\bar{v}$ and s from Sjögren and Svedberg (49). s is the average of five experiments with extremes of 3.41 and 3.55. D from Polson (44), and represents the average of three experiments.

LACTOGLOBULIN.—$\bar{v}$ and s from Pedersen (38). D from Lamm and Polson (23), Polson (44), and values of Polson given by Pedersen (38). Both s and D are functions of pH, but given constant molecular weights (s/D = constant) from pH 3 to 9.5. At constant pH the measurements of s vary about $\pm 3\%$, and of D $\pm 1\%$.

CHYMOTRYPSIN.—D from Kunitz and Northrop (22), and was obtained by the porous disc method. The value reported here was

obtained by multiplying their corrected values by 1.11, a factor derived by a careful examination of the porous disc results of McBain (25), Mehl (28), Polson (43) and Anson and Northrop (5). (We are indebted to Dr. J. W. Mehl for aid in this examination.)

CHYMOTRYPSINOGEN.—D and s are preliminary results of Hess and Williams (19). $\bar{v}$ is assumed to be 0.75.

PEPSIN.—D, s and $\bar{v}$ are largely from work of Polson (44) and Eriksson-Quensel (13). Philpot and Eriksson-Quensel (41) gave $s = 3.3$ (mean of seven determinations with estimated standard deviation of 0.15). Herzog (18) gave $D = 9.19$ in 1907.

EGG ALBUMIN.—D from Lamm and Polson (23) and Polson (44). Also in good agreement with the value 7.7 given by Tiselius and Gross (62). $\bar{v}$ from Nichols (32) and Svedberg and Nichols (55). s from Pedersen (39). Sjögren and Svedberg (48) gave $s = 3.54$, Svedberg and Eriksson (53) gave 3.4, Williams and Watson (63) gave 3.6 and Pickels (42) found 3.5. Earlier measurements of Svedberg and Nichols (55) and Nichols (32) gave 3.31, and indicated difficulties in the proper evaluation of the molecular weight by sedimentation velocity and diffusion data. Herzog (18) gave 7.54 for D in 1907.

SERUM ALBUMIN.—D from Polson (44), the measurements varying by ± 20%. Also from Kekwick (20). A value of 6.45 was given by Lamm and Polson (23). s from Mutzenbecher (30) and from Kekwick (20). $\bar{v}$ is from Svedberg and Sjögren (58) who found $s = 4.4$ and $D = 5.8$ (measured in ultracentrifuge by old technique). Two different serum albumin fractions of McMeekin (27) gave $s = 4.1$.

HEMOGLOBIN.—$\bar{v}$ from Svedberg and Fåhraeus (54). s from Pedersen and Andersson (40). A value of 4.63 was given by Steinhardt (52). Svedberg and Nichols (56) gave $s = 4.5$. Polson (44) and Lamm and Polson (23) gave $D = 6.9$ from measurements of human hemoglobin, and values for that of the horse a few per cent higher. The value 6.3 (57) was obtained by Tiselius and Gross (62), but would appear from these newer results to be too small. We have used $D = 6.9$ in our calculations. A value of 6.0 was also given by Svedberg and Nichols (56) as measured in the ultracentrifuge.

EDESTIN.—s and $\bar{v}$ from Svedberg and Stamm (60). Preliminary measurements by Oncley (35) give a somewhat larger value of $s = 14.6$. D is from Polson (44). Using the old method of measurement in the ultracentrifuge, Svedberg and Stamm (60) gave $D = 5.6$.

EXCELSIN.—D from Polson (44). s is the revised value (57) obtained from the experiments of Svedberg and Sjögren (59), who reported 11.8 as the average of nine experiments. $\bar{v}$ was obtained by these workers.

BUSHY-STUNT VIRUS.—s and $\bar{v}$ from McFarlane and Kekwick (26). $\bar{v}$ was the mean of two values (0.743 and 0.723), and s of eleven values varying from 138 to 158.

TOBACCO SEED GLOBULIN.—A single value of 12.7 for the sedimentation velocity constant of tobacco seed globulin was measured by Philpot and published by Crowfoot and Fankuchen (12). They assumed that this molecule was similar to edestin, excelsin, and amandin, and took an approximate molecular weight of 300,000.

Relaxation Times from Dielectric Dispersion

INSULIN.—From Cohn, Ferry, Livingood and Blanchard (11). This value was obtained in propylene glycol solutions with from 0 to 20% water present, and was corrected to values for water at 25°. The values obtained with the various water contents all gave the same value for the corrected relaxation time ($\pm$ 10%) in spite of a 100% change in viscosity. The value obtained for the relaxation time is too small to be explained on the basis of a molecular weight of 35,000–40,000, however, and further work is in progress on this problem.

LACTOGLOBULIN.—Based on measurements of Ferry, Cohn, Oncley and Blanchard (15), of Ferry and Oncley (17), and of Shack (47) (by the calorimetric method). See also Oncley, Ferry and Shack (37).

EGG ALBUMIN.—Measurements of Oncley (36) and Shack (47). See Oncley, Ferry and Shack (37).

SERUM ALBUMIN.—Measurements by Oncley (35) on serum albumin fractions prepared by McMeekin (27). Ferry and Oncley (16) gave approximately the same results on less highly purified fractions.

HEMOGLOBIN.—Measurements on horse hemoglobin by Oncley (34) and by Shack (47). See also Oncley, Ferry and Shack (37). A value of 17×10^{-8} for the relaxation time of pig hemoglobin was obtained by Arrhenius (6).

EDESTIN.—Measurements of Oncley (35). The results were complicated by the presence of about 10% of a higher molecular weight substance.

Viscosity

LACTOGLOBULIN.—From Polson (45). The data extends from 1 to 5% protein concentration.

PEPSIN.—From Polson (45). The concentration range was from 1 to 3%

EGG ALBUMIN.—Polson (45) gave 5.8, while measurements of Mehl, Blanchard and Walker (29) gave the somewhat lower result of 5.2. The value in TABLE 1 is an average of these results.

SERUM ALBUMIN.—Obtained by consideration of results published by Fahey and Green (14), Neurath and Saum (31) and a single point of Polson (45).

HEMOGLOBIN.—Polson (45) gave 5.3, while measurements of Cohn and Prentiss (10) and of Kunitz, Anson and Northrop (21) lead to somewhat lower results. The value given in TABLE 1 is an average of these results.

References for Notes on Table 1

1 Adair, G. S. Skand Arch. Physiol. **49**: 76. 1926.
2 Adair, G. S. Proc. Roy. Soc. London A**108**: 627. 1926; **120**: 573. 1928.
3 Adair, G. S. Ann. Rev. Biochem **6**: 180. 1937; Jour. Am. Chem. Soc. **49**: 2524. 1927.
4 Adair, G. S., & Robinson, M. E. Biochem. Jour. **24**: 1864. 1930.
5 Anson, M. L., & Northrop, J. H. Jour. Gen. Physiol. **20**: 575. 1937.
6 Arrhenius, S. Physik. Zeit. **39**: 559. 1938.
7 Burk, N. F. Jour. Biol. Chem. **98**: 353. 1932.
8 Burk, N. F. Jour. Biol. Chem. **120**: 63. 1937.
9 Burk, N. F., & Greenberg, D. M. Jour. Biol. Chem. **87**: 197. 1930.
10 Cohn, E. J., & Prentiss, A. M. Jour. Gen. Physiol. **8**: 619. 1927.
11 Cohn, E. J., Ferry, J. D., Livingood, J. J., & Blanchard, M. H. Science **90**: 183. 1939.
12 Crowfoot, D., & Fankuchen, I. Nature, **141**: 521. 1938.
13 Eriksson-Quensel, I.-B. unpublished work.
14 Fahey, K., & Green, A. A. Jour. Am. Chem. Soc. **60**: 3039. 1938.
15 Ferry, J. D., Cohn, E. J., Oncley, J. L., & Blanchard, M. H. Jour. Biol. Chem. **128**: Proc. 28. 1939.
16 Ferry, J. D., & Oncley, J. L. Jour. Am. Chem. Soc. **60**: 1123. 1938.
17 Ferry, J. D., & Oncley, J. L. Jour. Am. Chem. Soc. **63**: 272. 1941.
18 Herzog, R. O. Zeit. Elektrochem. **13**: 533. 1907.
19 Hess, E. E., & Williams, J. W. unpublished work.
20 Kekwick, R. A. Biochem. Jour. **32**: 552. 1938.
21 Kunitz, M., Anson, M. L., & Northrop, J. H. Jour. Gen. Physiol. **17**: 365. 1934.
22 Kunitz, M., & Northrop, J. H. Jour. Gen. Physiol. **18**: 433. 1935; **20**: 575. 1937.
23 Lamm, O., & Polson, A. G. Biochem. Jour. **30**: 528. 1936.
24 Marrack, J., & Hewitt, L. F. Biochem. Jour. **23**: 1079. 1929.
25 McBain, J. W., & Liu, T. H. Jour. Am. Chem. Soc. **53**: 59. 1931; McBain, J. W., & Dawson, C. R. Proc. Roy Soc. London A**148**: 32. 1935.
26 McFarlane, A. S., & Kekwick, R. A. Biochem. Jour. **32**: 1607. 1938.
27 McMeekin, T. L. Jour. Am. Chem. Soc. **61**: 2884. 1939; **62**: 3393. 1940.
28 Mehl, J. W., & Schmidt, C. L. A. Univ. Cal. Publ. Physiol. **8**: 165. 1937.
29 Mehl, J. W., Blanchard, M. H., & Walker, P. H. unpublished work.
30 Mutzenbecher, P. v. Biochem. Zeit. **266**: 226, 250, 259. 1933.
31 Neurath, H., & Saum, A. M. Jour. Biol. Chem. **128**: 347. 1939.
32 Nichols, J. B. Jour. Am. Chem. Soc. **52**: 5176. 1930.
33 Northrop, J. H. Jour. Gen. Physiol. **13**: 739. 1930.
34 Oncley, J. L. Jour. Am. Chem. Soc. **60**: 1115. 1938.
35 Oncley, J. L. Jour. Phys. Chem. **44**: 1103. 1940.
36 Oncley, J. L. unpublished work.
37 Oncley, J. L. Ferry, J. D., & Shack, J. Ann. N. Y. Acad. Sci. **40**: 371. 1940.
38 Pedersen, K. O. Biochem. Jour. **30**: 961. 1936.
39 Pedersen, K. O. unpublished work.
40 Pedersen, K. O., & Anderson, K. J. I. unpublished work.
41 Philpot, J. St. L., & Eriksson-Quensel, I.-B. Nature **132**: 932. 1933. Philpot, J. St. L. Biochem. Jour. **29**: 2458. 1935.

42 Pickels, E. G. unpublished work.
43 Polson, A. G. Biochem. Jour. 31: 1903. 1937.
44 Polson, A. G. Kolloid. Zeit. 87: 149. 1939
45 Polson, A. G. Kolloid. Zeit. 88: 51. 1939.
46 Roche, J., Dorier, M., & Marquet, F. Compt. rend. Soc. Biol. 119: 1150. 1935.
47 Shack, J. Ph.D. Dissertation, Harvard University. 1939.
48 Sjögren, B., & Svedberg, T. Jour. Am. Chem. Soc. 52: 5187. 1930.
49 Sjögren, B., & Svedberg, T. Jour Am. Chem. Soc 53: 2657. 1931.
50 Sörensen, S. P. L. Compt. rend. trav. lab. Carlsberg 12: 262. 1917.
51 Sörensen, S. P. L. "Proteins." New York. 1925.
52 Steinhardt, J. Jour. Biol. Chem. 123: 543. 1938.
53 Svedberg, T., & Eriksson, I.-B. Jour. Am. Chem. Soc. 56: 409. 1934.
54 Svedberg, T., & Fåhraeus, R. Jour. Am. Chem. Soc. 48: 430. 1926.
55 Svedberg, T., & Nichols, J. B. Jour. Am. Chem. Soc. 48: 3081. 1926.
56 Svedberg, T., & Nichols, J. B. Jour. Am. Chem. Soc. 49: 2920. 1927.
57 Svedberg, T., & Pedersen, K. O. "The Ultracentrifuge." Oxford Univ. Press
    London. 1940
58 Svedberg, T., & Sjögren, B. Jour. Am. Chem. Soc. 50: 3318. 1928.
59 Svedberg, T., & Sjögren, B. Jour. Am. Chem. Soc. 52: 279. 1930.
60 Svedberg, T., & Stamm, A. J. Jour. Am. Chem. Soc. 51: 2170. 1929.
61 Taylor, G. L., Adair, G. S., & Adair, M. E. Jour. Hygiene 32: 340. 1932
62 Tiselius, A., & Gross, D. Kolloid. Zeit. 66: 11. 1934.
63 Williams, J. W., & Watson, C. C. Nature 139: 506. 1937.

THE X-RAY DIFFRACTION METHODS USED IN PROTEIN STUDIES

By B. E. Warren

From the George Eastman Laboratory of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts

The following discussion is intended primarily for those who are not specialists in X-ray diffraction methods, and is planned to give an elementary picture of the ideas involved, the methods used, the kind of information obtained, and the reliability of the conclusions which can be drawn.

A crystal is a three-dimensional scheme of repetition, built up by some unit of structure repeating itself identically at regular intervals in three dimensions. The scheme of repetition is specified by three axes, a, b, c, which give the shortest translations between identical points. FIGURE 2a indicates schematically in two dimensions the repeating nature of a crystalline structure. The volume of the parallelopipedon defined by the three axes, a, b, c, is called the unit cell. This is the smallest volume of crystal which repeated can produce the crystal as a whole. The problem of crystal structure analysis divides into two stages: the determination of the size and shape of the unit cell (axes a, b, c, and angles); and the positions of all the atoms within the unit cell.

For complex crystals such as proteins, the single crystal rotation method is generally used. The method is indicated schematically in FIGURE 1. A single crystal is mounted so that it can be rotated about one of its crystallographic axes, say c. A monochromatic X-ray beam (usually the $K\alpha$ line of copper, $\lambda = 1.539\text{Å}$) is collimated by passing through two pin holes. The X-ray beam is scattered by every atom in the crystal, and the secondary waves coming from the atoms will show constructive interference in certain directions, and will cancel out in other directions. A series of diffracted beams will result, and these beams are recorded on either a cylindrical film or a flat film. If the experimental arrangement is as indicated in FIGURE 1, the diffracted beams will all lie on a set of cones, which intersect a cylindrical film to give a series of layer lines.

Measuring the layer line separation y_1 one computes the angle β_1 by the relation

$$\tan \beta_1 = \frac{y_1}{R}$$

The length of the crystallographic axis about which the crystal was rotated is then given by

$$c = \frac{\lambda}{\sin \beta_1} \qquad (1)$$

Knowing the wave length λ, and measuring the layer line separation, one gets directly the length of the axis. The other two axial lengths are then obtained by rotation patterns rotating around the corresponding axes. The only possibility of mistake comes from overlooking any very weak layer lines, but this is not likely to happen often.
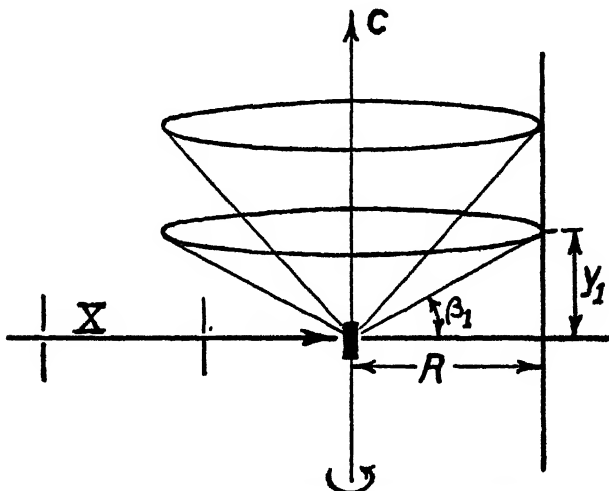

FIGURE 1. Schematic arrangement of X-ray beam and crystal in making rotation patterns of single crystals.

The dimensions of the unit cell are therefore obtained directly from the geometry of the diffraction pattern, and in general can be taken as absolutely reliable.

If n is the number of molecules per unit cell, M the molecular weight, N the Avogadro number, and ρ the density of the crystal, expressing the mass of the unit cell in two different ways gives the following relation:

$$\frac{nM}{N} = V\rho \qquad (2)$$

$$n = \frac{NV}{M}\rho \qquad (3)$$

$$M = \frac{NV}{n}\rho \qquad (4)$$

Structure
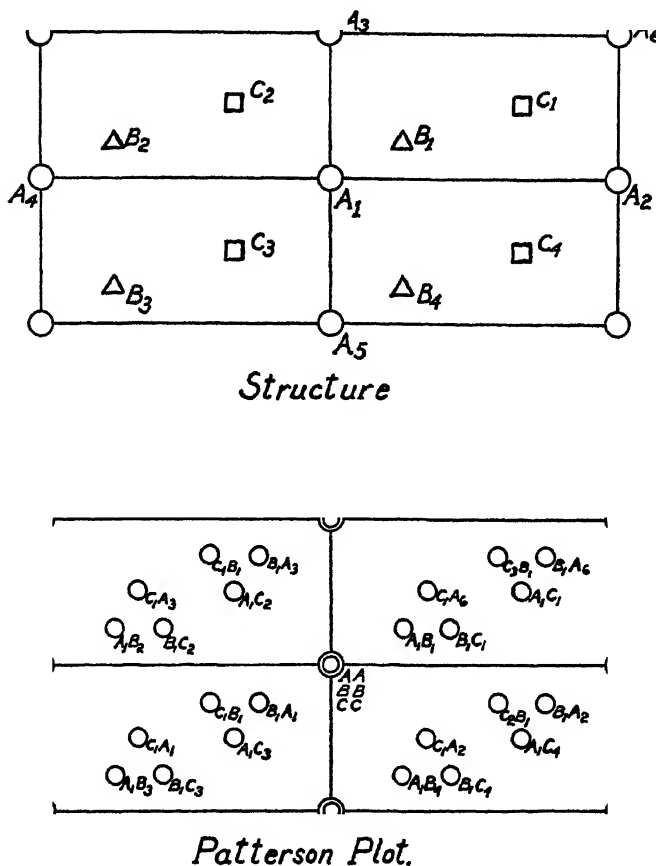
Patterson Plot.

FIGURE 2. Schematic representation in two dimensions of a crystalline structure and the corresponding Patterson plot.

If one knows the molecular weight M, the number of molecules per unit cell can be calculated from (3). If one knows the number of molecules per cell, the molecular weight can be calculated from (4). In general M cannot be obtained uniquely from X-ray data alone. Using (4) one can calculate a series of possible molecular weights to be compared with determinations by other methods. Often the lattice and the symmetry of the crystal will impose certain restrictions upon n and consequently upon M. The X-ray value of the molecular weight of a protein means one of the series of values calculated by (4) which agrees best with ultracentrifuge values, and involves a value of n which is compatible with the symmetry of the crystal. The

numerical value of M is uncertain to the same extent as the value of the density which is used in the computation.

The next step in the structure determination involves finding the position of each atom in the unit cell. This is by far the most difficult step, there is no straightforward method, and often it has turned out to be too difficult to be carried through. The positions of the atoms determine the intensities of the diffracted beams, and atomic positions must be deduced from intensity considerations.

The use of symmetry considerations in determining atomic positions is often of great help. This kind of reasoning is known as application of space group theory. It turns out that there are 230 possible combinations of symmetry operations which can exist in a crystalline structure. The statement of the space group of a crystal is simply the statement of the kind of symmetry operations and their spacial arrangement.

The structure factor, F, is the amplitude of scattered radiation per unit cell. It involves the positions of the atoms in the unit cell (coordinates x, y, z) and is given by the expression

$$F_{hkl} = \sum_n f_n \, e^{2\pi i\left(\frac{hx_n}{a} + \frac{ky_n}{b} + \frac{lz_n}{c}\right)} \tag{5}$$

h, k, l are the Miller indices of the diffracting planes, f_n is the amplitude of scattered radiation from atom n. What one observes experimentally is the integrated intensity of a diffracted beam, and this quantity is given by the equation

$$I_{hkl} = K \, F_{hkl}{}^2 \tag{6}$$

The factor K is known for any experimental arrangement and diffracted beam hkl, so that from the intensities of the diffracted beams one gets directly a series of values for $F_{hkl}{}^2$.

Right at this point comes the stumbling block of X-ray crystallography. Since it is F^2 which is determined experimentally, one gets only the magnitude of F without knowing the phase. For example, in a simple case one does not know from measurements of F^2 whether F is positive or negative. If one could measure F directly it would be possible by Fourier series methods to plot out the distribution of scattering matter in the unit cell, and hence to determine the atomic positions directly from the X-ray data. However, this is impossible.

We are forced to carry on by indirect methods which fall into three classes.

(a) Cut and try methods. Using all available physical, chemical, and symmetry data to guess reasonable structures, and then comparing the calculated F^2 values with those obtained from the intensities. Notice the one-way nature of the equations: from a structure one can calculate all the F^2 values, but from the F^2 values one cannot directly calculate a structure.

(b) Use of heavy atoms, either in the structure, or put in by isomorphous replacement. Sometimes this allows the positions of the heavy atoms to be determined independently; sometimes it is used to determine the phase of reflections; sometimes it is used in conection with Patterson plots.

(c) The method of F^2 series, Patterson or Patterson-Harker plots. In case no information is available other than the X-ray intensity data, the only method of making progress involves using the Patterson type Fourier plots. A Patterson plot is given by a Fourier series in which the coefficients are the F_{hkl}^2 quantities, which themselves are determined directly from the experimental intensities.

$$V(X, Y, Z) = \sum \sum_{-\infty}^{+\infty} \sum F_{hkl}^2 \cos 2\pi\left(\frac{hX}{a} + \frac{kY}{b} + \frac{lZ}{c}\right) \tag{7}$$

The Patterson 2-dimensional projection is given by the equation

$$V(X, Y) = \sum\sum_{-\infty}^{+\infty} F_{hko}^2 \cos 2\pi\left(\frac{hX}{a} + \frac{kY}{b}\right) \tag{8}$$

Since the F_{hkl}^2 values are determined directly by experiment, it is evident that the Patterson type series can be evaluated and plotted out. It should be emphasized that a Patterson plot does not directly represent a structure.

The significance of a Patterson plot is illustrated in Figure 2. The upper figure shows four unit cells of a crystal containing three atoms per unit cell. In the lower figure the position of each peak with respect to the origin gives the position of some atom in the structure with respect to another atom. For example, the displacement of the peak B_1C_1 from the origin of the plot represents the displacement of the atom C_1 from atom B_1. The peaks of a Patterson plot give the positions of each atom with respect to every other atom, and not the positions of the atoms with respect to a single origin.

It is obvious that for a complex structure the Patterson plot may be very involved and the interpretation extremely difficult. Whereas in theory it should be possible to work out quite complex structures

from a Patterson plot, in actual practice it is far more difficult than would be imagined.   The difficulties are mostly of a practical nature: ghost peaks due to breaking off the series, effects of errors in the experimental intensities, and, worst of all, superposition and overlapping of the peaks.

EVIDENCE FROM X-RAYS REGARDING THE STRUCTURE OF PROTEIN MOLECULES

By I. Fankuchen*

From Massachusetts Institute of Technology, Cambridge, Massachusetts

The problem of the internal structure of the protein molecule is one on which very little real progress has been made since the fundamental researches of Fischer. His work showed that proteins were composed of *a*-amino acid residues bound together by peptide links. Other papers presented at this conference have dealt in detail with the chemical evidence on the constitution of proteins.

In recent years, the application of X-ray diffraction methods to the study of protein structures has offered a new approach to this very fundamental problem. Astbury's[1, 2] work has given us much information as to the stereochemical configuration of the fibrous proteins. These contain polypeptide chains, fully extended as in silk or folded in various ways as in wool. The detailed nature of the folding, however, has not been established even in these comparatively simple proteins. The crystalline proteins, *when denatured*, also appear to contain polypeptide chains.[3] More recently, the use of X-ray methods has been extended to the study of the native crystalline proteins. The diffraction effects are quite complex and at certain stages of the analysis, differences of opinion concerning the interpretation of the results have arisen. Enough work has by now been published to make a critical examination of the techniques and data desirable. Such a critical examination should show where assumptions enter and by making clearer the distinction between fact and theory help put the X-ray study of proteins on a sounder basis.

The first X-ray studies in which diffraction effects were clearly obtained all used the powder technique. The specimen consists of a poly-crystalline mass of material irradiated by a comparatively mono-chromatic X-ray beam. By the use of this method, squash seed globulin,[4] chymotrypsinogen,[5] the Bence-Jones protein,[6] urease and

* National Research Fellow in Protein Chemistry, Massachusetts Institute of Technology.
1 Astbury, W. T. Jour. Textile Inst. 27: 281. 1936.
2 Astbury, W. T. Science Progress No. 133. July 1939.
3 Astbury, W. T. Nature 140: 968. 1937.
4 Astbury, W. T., & Lomax, R. Jour. Chem. Soc. 846. 1935.
5 Corey, R. B., & Wyckoff, R. W. G. Jour. Biol. Chem. 114: 407. 1936.
6 Magnus-Levy, A., Meyer, K. H., & Lotmar, W. Nature 137: 616. 1936.

pepsin,[7] egg albumin,[8] and hemoglobin,[9, 10] were all shown to give sharp diffraction rings. These observations were taken as proving the true crystallinity of the materials studied. This conclusion, while undoubtedly correct in the above cases, does not necessarily follow from the observation of sharp diffraction rings. The possibility of error arises from the great size of the molecules. It is true that the production of sharp diffraction effects indicates a periodic structure, but such a structure may well exist within the molecule itself and the rings may therefore be no proof of the crystalline arrangement of the molecules relative to one another. Such a situation has been found to be the case for tobacco mosaic virus.[11] Of course, if the particles are large enough to be studied in the microscope and are there seen to be *truly* crystalline, such considerations need not disturb us. Except for simple cases like cubic and hexagonal close-packed systems, powder data from proteins by themselves are not easy to interpret. One can compute the spacings responsible for the diffraction rings and perhaps if such questions arise, settle questions of identity, but it is difficult to go much further. For this reason, the data from the above researches will not in general be considered in the discussion to follow and are not included in the tabulated data.

The study of protein structures by single crystal methods has, up to now, been done primarily by Bernal and his school. TABLE 1, in addition to the results of single crystal studies also includes two proteins, tobacco seed globulin and the bushy stunt virus of tomato, which were studied by the powder method. The symmetry of these two proteins, both cubic, permitted unit cell and molecular weight determinations to be made and they were therefore included. The unit cells for insulin and excelsin are referred to orthohexagonal axes. Their primitive cells are rhombohedral and contain only one molecule per unit cell.

The values of the unit cell dimensions are obtained from measurements of the *locations* of the X-ray reflections and therefore constitute the surest quantitative information we can derive from these X-ray studies. For ordinary crystals the accuracy of spacing measurements can be pushed as high as 1 in 100,000 but for protein crystals nothing like this accuracy can be hoped for. The small size of most protein crystals, the magnitude of the spacings (and consequent small angles of deviation of the X-ray beam), the imperfections of the crystals and

[7] Fankuchen, I.   Jour. Am. Chem. Soc. **56**: 2398.   1934
[8] Clark, G. L., & Shenk, J. H.   Radiology **28**: 53.   1937.
[9] Clark, G. L., & Shenk, J. H.   Radiology **28**: 144.   1937.
[10] Wyckoff, R. W. G., & Corey, R. B.   Science **81**: 365.   1935.
[11] Bawden, F. C., Pirie, N. W., Bernal, J. D., & Fankuchen, I.   Nature **138**: 1051.   1936.

the low intensity of the reflections all combine to make accurate measurements difficult. With reasonable precautions the probable error can be reduced to one part in 250 and consequently the volume of the unit cell can be found to within approximately 1%. For present needs this accuracy is adequate. Future developments may require higher accuracy in spacing determination and it is probable that the accuracy of the spacing determination can be increased to a limit set by the inherent imperfections of the periodic structure of the crystals.

If the volume of the unit cell is known, the determination of the density enables the direct computation of the mass of the unit cell to be made. The mass of the hydrogen atom (atomic weight 1.008) is 1.66×10^{-24} grams. Thus the molecular weight of the cell contents can be obtained by dividing the mass of the unit cell by 1.65×10^{-24}. If the number, n, of molecules in the cell is known, a simple division gives the molecular weight of the protein molecule. A number of questions may now well be asked. What densities are to be used and how are they to be measured? How is n, the number of molecules in the unit cell, determined? In the attempt to answer these questions, we become aware of some serious difficulties.

Single protein crystals have been studied both dry and immersed in solutions and wherever both wet and dry crystals of the same protein have been used, appreciable differences in the unit cells occur (see TABLE 1). In computing molecular weights, the density to be used is the density of the crystal in the medium in which it is studied—for dry crystals the density in air and for wet crystals, the density in the solution. However, to measure densities, immersion in some liquid is usually necessary. It has been shown by several workers[12, 13] that appreciable differences in the densities are found depending on the medium used to immerse the crystals. It has also been shown[14] that variations in the cell dimensions exist when the crystals are studied in different solutions. The molecular weights obtained from centrifuge data are presumed to correspond to anhydrous molecular weights[15] and are therefore to be compared directly with the molecular weights determined from X-ray measurements on dry crystals (values corrected for residual water where possible). But this comparison is meaningful only if the density used in the X-ray determination is accurately the density of the dry crystal. Similar difficulties occur for the wet crystals. These give the best X-ray diffraction effects and conse-

[12] Adair, G. S., & Adair, M. E. Proc. Roy. Soc. London **B120: 422.** 1936.
[13] Crowfoot, D. Proc. Roy. Soc. London **A164: 580.** 1938.
[14] Crowfoot, D., & Riley, D. Nature **144: 1011.** 1939.
[15] Svedberg, T. Proc. Roy. Soc. London **B127: 7.** 1939.

quently a knowledge of the accurate X-ray wet molecular weights would be most useful if we knew precisely what they meant. Thus the dry molecular weights could be computed if we knew the state of hydration of the wet crystals. Until we know more about the densities of protein crystals and the role of the ions in the solutions in which they are immersed, any comparisons should be made very cautiously.

The determination of the number, n, of molecules in the unit cell always involves the introduction of non X-ray data. The number must be an integer, usually a small one, and the symmetry of the space group imposes further restrictions on the possible values of n. Dividing the cell molecular weight by the possible values of n gives a set of possible molecular weights, the maximum of which is the cell molecular weight. One then chooses such a value for n from the set of possible values as will give a molecular weight in best agreement with the values given by other techniques: ultracentrifuge, chemical analysis, etc. In general, a very rough idea of the molecular weight is sufficient to define n. This procedure sometimes imposes restrictions on the molecular symmetry. Thus in insulin, where there is one molecule in the unit cell, the molecules must have trigonal symmetry to within the accuracy of the X-ray data; and in horse methemoglobin, the choice of the molecular weight makes $n = 2$, and thereby imposes a twofold axis of symmetry on the molecule. The question may be purely academic, but it does seem possible that in some protein crystals the unit may be different from what it appears to be in the ultracentrifuge, let us say. In such a case, if we make use of the ultracentrifuge value in determining n, some erroneous deductions as to the size and symmetry of the protein unit in the crystal will then result.

Some further rough idea as to the size and shape of the molecules may be obtained from a knowledge of the unit cell size and space group. This about exhausts the information one can obtain from a study of the location of the reflections without considering their relative intensities. The choice of the space group depends to a large extent on a study of systematic absences of reflections—a special case of intensity of reflection zero—but such determinations of systematic absences are fairly free from pitfalls or ambiguities.

Such a statement unfortunately cannot be made of the extension of the analysis of the data to a study of the relative intensities of the reflections. The relative intensities of the various reflections are determined by the distribution of scattering matter within the unit cell and it is the ultimate goal of the structure analysis to work backwards from the intensities to the distribution. In the simplest crystals, the

X-ray data alone will suffice to determine the structure but it seems a fairly general rule that as the structures grow more complicated, an increasing amount of outside information is necessary. (In special cases[16] this information may consist of X-ray data on isomorphous crystals, and the suggestion has been made[17] that this method be applied to protein X-ray studies.) At the same time the way the X-ray data are used changes. Thus, in the simpler structures, one may work either from the X-ray data to the structure without making use of any preconceptions as to the answer, or reverse the procedure by assuming a structure (enough is now known of crystal chemistry to enable shrewd and often accurate guesses to be made), and then testing the assumption by computing what its X-ray reflections should be. As the structures grow more complicated, the second method tends to be used increasingly, and when structures as complex as the proteins are reached, one is sorely tempted to use the second method exclusively. Given a set of X-ray intensities, we can assume a structure based on some theory and then test our theory by a comparison of observed and computed intensities. For proteins, however, this testing must be done very carefully with one eye on the character of the data and another on the theory to be tested. Thus, if the X-ray data are limited in scope (as is often the case for proteins) agreement between observed and computed intensities can say nothing about any fine structure of the theory and indeed need not even be a confirmation of the large-scale details of the structure being tested.[18, 19]

Such questions are perhaps more easily discussed in terms of Fourier series and Patterson diagrams. Sir William Bragg showed[20] that the three-dimensional periodic structure of scattering matter in a crystal could be expressed as a three-dimensional Fourier series, the coefficients of which are the structure factors, F, of the various X-ray reflections. Unfortunately, we do not observe F directly. The X-ray intensity is proportional to $|F|^2$ and in general F may be a complex number whose amplitude is determinable from the X-ray intensities but whose phase is not. Various methods have been used for determining these unknown phases but so far they have not been applied to protein data.

Patterson[21, 22] considered the Fourier series in which values of $|F|^2$ were used as coefficients instead of values of F, and showed that such

16 **Robertson, J. M.** Jour. Chem. Soc. 615. 1935; 1195. 1936.
17 **Robertson, J. M.** Nature 143: 75. 1939.
18 **Bragg, W. L.** Nature 143: 73. 1939.
19 **Bernal, J. D.** Nature 143: 74. 1939.
20 **Bragg, W. H.** Phil. Trans. A215: 253. 1915.
21 **Patterson, A. L.** Zeit. Krist. 90: 517. 1935.
22 **Patterson, A. L.** Zeit. Krist. 90: 543. 1935.

a series could be given a physical meaning. The three-dimensional summation of $|F|^2$ terms gives a periodic assemblage of peaks which represent the ends of interatomic vectors drawn from a common origin. Analagous to the case in the F series, two- and one-dimensional summations may be made which use only part of the available X-ray data and represent projections of the three-dimensional peak system onto a plane or line, respectively. Such two-dimensional projections have been most widely used for two reasons. The amount of data required is much less than for the three-dimensional summation, and the task of computing the series is also not so involved. The great drawback of the two-dimensional projection is the superposition of peaks which makes the unravelling of the pattern a most difficult and often impossible task. Harker[23] has pointed out that usually only certain sections of the three-dimensional summations are important. The computation then resolves itself into the summation of two- and one-dimensional series although for such sections the complete set of X-ray intensities are used. Robertson[24] has recently reviewed this field in a most comprehensive and clear paper.

Theoretically, the Patterson peaks can be related to interatomic vectors, and a knowledge of the contents of the unit cell should enable one in most cases to determine an unique periodic distribution of scattering matter which would correspond to the experimentally determined Patterson diagram. In the case of protein crystals, superposition of vectors occurs and the few peaks found in protein Patterson diagrams are the result of much collaboration between vectors. The Patterson diagrams can then test only gross features of the proposed structure. Indeed, while the Patterson summation contains all the X-ray data and theoretically an harmonic analysis of the summation would give the original set of F^2 values uniquely, yet the Patterson diagram, because it is a summation, conceals the fact that scores of individual intensities went into its making. It may be the best way of presenting the totality of X-ray data, but that does not necessarily make it the best way of using the data to test proposed structures. It may, for this purpose, be desirable to make use of the individual intensities. Here, absolute intensities would be most valuable. In any event, a limit is set by the minimum spacings observed on the fine grain structure which can be tested.

It would appear desirable, from this discussion, to work if possible from the X-ray data to a structure instead of attempting to use the

[23] Harker, D.   Jour. Chem. Phys. 4: 381.   1936.
[24] Robertson, J. M.   Rept. Progress Phys. Physical Soc. 4: 332.   1937.

X-ray data to test structures. The rate of progress may then be painfully slow, but to compensate, such a procedure reduces the danger of believing that more has been proved than is actually the case. The whole problem is a difficult one. However, enough progress has already been made so that one may justly believe that a slow, cautious approach will materially contribute to a fuller knowledge of the structure of proteins.

An attempt will now be made to summarize the information we have obtained about protein structures as a result of X-ray studies. The cell size and molecular weight determinations are listed in TABLE 1. A survey of the character of the X-ray data gives some information as to the regularity of the structures. Uniformly, the crystals immersed in solutions give far better photographs than the dry crystals, the reflections being more numerous and intense, better defined and what is most significant, reaching out to much smaller spacings than is the case for the dry crystals. This suggests not only some disorientation of the molecules in the dry crystals, but probably also some minor rearrangement within the molecules. In all cases where both wet and dry crystals of the same protein have been studied, cell shrinkages have been observed. It is reasonable to assume that upon drying most of the water is removed from between the molecules which then rearrange themselves while remaining substantially unchanged as molecules. It has been suggested[27] that by studying the changes in the Patterson diagrams as we pass from wet to dry proteins, one will be able to separate the effects due to inter- and intramolecular scattering.

Perhaps the most important characteristic of the intensity patterns is the general enhancement of the reflections in the neighborhoods of 10Å and 4.5Å for wet crystals and 10Å for the dry ones.[27, 82] These dimensions were found by Astbury[1] to be the periodicities of the packing of the polypeptide chains in directions at right angles to their length, and correspond to the average length of the side chains and backbone spacing, respectively. It is thought that the change in going from a crystalline protein to the denatured form, while fundamental, does not involve any major rearrangement of the constituent

[23] Crowfoot, D. Nature **135**: 591. 1935.
[24] Crowfoot, D., & Riley, D. Nature **141**: 521. 1938.
[27] Bernal, J. D., Fankuchen, I., & Perutz, M. Nature **141**: 523. 1938.
[28] Bernal, J. D., & Crowfoot, D. Nature **133**: 794. 1934.
[29] Crowfoot, D., & Fankuchen, I. Nature **141**: 522. 1938.
[30] Astbury, W. T., & Bell, F. O. Tab. Biol. 17 (1): 106. 1939.
[81] Bernal, J. D., Fankuchen, I., & Riley, D. Nature **142**: 1075. 1938.
[82] Bernal, J. D. Nature **143**: 663. 1939.

TABLE 1

Protein and References	Insulin (18, 14, 25)	Lactoglobulin (26) (46) Tabular	Lactoglobulin (26) (46) Needle	Chymo-Trypsin (27) (45)	Pepsin (28)	Horse Met-hemoglobin (27) (45)	Tobacco Seed Globulin (29)	Excelsin (30) (33)	Bushy Stunt Virus (31)
Non X-ray Molecular Weight	35,100 / 40,900	37,900 / 41,800		41,000	35,500–39,200	66,700	300,000	294,000	7,600,000 / 8,800,000
a in Å. (10^{-8} cm.)	130	60	56	45		102	123	149	318
b	74.8	63	56	62.5		51	123	86	318
c	30.9	110	130	57.5		47	123	208.2	318
β	90°	90°	90°	112°		130°	90°	90°	90°
c sin β	30.9	110	130	53.5		36	123	208.2	318
Vol. in Å³	298,000	416,000	408,000	151,000		188,000	1,850,000	2,670,000	32,000,000
n	6	8	8	2		2	4	6	2
A Volume per Molecule	50,000	52,000	51,000	75,500		94,000	460,000	445,000	16,000,000
M Density	1.315	1.27	1.30	1.31		1.270	1.287	1.317	1.35
R Molecular Weight	39,500	40,000	40,100	60,000		72,000	360,000	350,000?	13,000,000
D Molecular Weight Corrected for Residual Water	37,400			54,000		66,700	322,000	305,800	
Space Group	R3	P2₁2₁2₁		P2₁		C2	F	R3	I
Smallest Observed Spacing	7	20	20	5		13			
	a =44.4, α =114°28′, n =1, Referred to rhombohedral unit cell							a =85.3, α =60°36′, n =1	
a	144	67.5	67.5	49.6	116	110			394
b	83	67.5	67.5	67.8	67	63.8			394
c	34	154	133.5	66.5	461	54.2			394
β	90°	90°	90°	102°	90°	112°			90°
F c sin β / Vol. in Å³	34 / 404,000	154 / 702,000	133.5 / 608,000	65 / 219,000	461 / 3,580,000	50.2 / 352,000			394 / 61,000,000
H n	6	8	8	2	54	2			2
Volume per Molecule	67,000	88,000	76,000	109,000	66,500	176,000			30,500,000
Wet Molecular Weight	52,400	67,000		84,500	53,000	132,000			24,000,000
Density	1.28	1.257	1.30	1.277	1.32	1.242			1.286
Space Group	R3	P2₁2₁2₁	P4₂2₁	P2₁		C2			I
M Smallest Observed Spacing	2.4	2.4		2		2			1
	a =49.4, α =114°16′, n =1								

X-RAY DATA

Insulin: Crowfoot[12][25] (dry), and **Crowfoot & Riley**[14] (wet).

Lactoglobulin: Crowfoot & Riley.[26] Further work is being carried on by these workers (private communication) including Fourier studies and X-ray studies of crystals in various states of hydration.[46]

Chymotrypsin: Bernal, Fankuchen & Perutz,[27] Perutz.[45] (In the original communication[27] n was taken as 4. It is now suggested by Perutz that n = 2. This gives a molecular weight in better agreement with the data.) Crystals prepared by Northrop.

Hemoglobin: Bernal, Fankuchen & Perutz.[27] Further intensive study is being done by Perutz (private communication).[46]

Pepsin: Bernal & Crowfoot.[18] These measurements were the first on single protein crystals and repetition of this work would be desirable.

Tobacco Seed Globulin: Crowfoot & Fankuchen.[19] Fairly old, dry crystals were used prepared by Vickery. Only three extremely faint powder lines were observed. It would be desirable to have these measurements repeated and an effort to obtain data on wet crystals would be worthwhile.

Excelsin: Astbury, Dickinson & Bailey[23] and Astbury & Bell (unpublished work). Astbury & Bell[20] refer to this unpublished work in their paper in Tabulae Biologicae.

Bushy Stunt Virus: Bernal, Fankuchen & Riley.[21] Material furnished by Bawden & Pirie.

NON X-RAY MOLECULAR WEIGHT DATA

For insulin; lactoglobulin, pepsin and excelsin are taken from Svedberg[15] and are centrifuge measurements.

Chymotrypsin: 41,000, Kunitz & Northrop,[44] measured by osmotic pressure.

Hemoglobin: 66,700 is found by chemical methods. The ultracentrifuge value for horse hemoglobin[15] is 63,000–69,000.

For **Tobacco Seed Globulin:** only the sedimentation constant 12.7×10^{-13} was known (Philpot, unpublished data). This is similar to the values for the other seed globulins whose molecular weights are about 300,000.

Bushy Stunt Virus: molecular weights 7,600,000 and 8,800,000, McFarlane & Kekwick.[41] These molecular weights were used in determining "n" the number of molecules per unit cell.

DENSITIES

Insulin: both needle-shaped and flat rhombohedra were studied by Crowfoot. These were shown to possess identical crystal structure. The needles were imperfect and gave density values a trifle lower than the rhombohedra. The value chosen was the highest observed and was obtained from the largest rhombohedral crystals. Residual water 5.35% in air-dried crystals was determined by drying under reduced pressure at 104°.

Lactoglobulin: Wet Tabular, 1.257 in sugar solution (Crowfoot).[26] Wet Needles, dissolved too rapidly to permit of any density determinations. Dry Crystals, found to be 1.27 by immersion in o-dichlorbenzene and toluene.[26] Crowfoot believes this value too low, due to occlusion of air and uses an assumed density of 1.31 (insulin) in computations of molecular weights.

Chymotrypsin: Wet. Determined by Perutz.[27] Dry, value of 1.31 assumed.[27]

Hemoglobin: Wet. Determined by Perutz.[27] Dry, value of 1.26 assumed to correspond with value for serum albumin as given by Chick and Martin.

Tobacco Seed Globulin: measured ni sodium phosphate buffer solution at pH 5.0. Residual water 10.4%. Determined by drying in vacuum at 100°.

Excelsin: Value of 1.31 assumed by the writer to permit computation of molecular weight. Table in paper by Astbury & Bell gives only unit cell data and dry molecular weight. Assumption is made that this value is corrected for residual water.

Bushy Stunt Virus: Wet crystals, 1.286, measured by immersion. Dry crystals, 1.35, from computations of McFarlane & Kekwick.[41]

atoms.[32, 33, 34] Thus a single crystal of excelsin upon partial denaturation[33] gives a composite fiber and single crystal photograph in which the fibers are oriented in the direction of the crystal axes. Undoubtedly, these observations are connected with something fundamental in protein architecture, but thus far no sure explanation has been advanced. It has been suggested[32] that in crystalline proteins an arrangement of side chains and backbone spacings occurs, similar to that found in fibrous proteins. This suggestion must be considered as purely tentative since enough confirmatory evidence is lacking.

The protein crystal which has been studied most intensively is insulin.[13, 14, 25] The dry crystal which until recently was the only form studied did not give reflections comparable in quantity or quality with those obtained from other protein crystals, but it possessed the great advantage of having only one molecule per unit cell. This fact resulted in a great simplification in the Patterson diagrams which were the first step in an attempted analysis of the X-ray data. Even then, the interpretation of these diagrams had led to acute controversy which it is not my intention to enter into here. The papers dealing with this subject are given below.[14, 17, 18, 19, 32, 34, 35, 36, 37, 38, 39, 40, 42, 43] *

It is desirable to see just how far the study of insulin has been carried by Crowfoot.[13] The unit cell determinations are quite straightforward and the cell volumes are correct to about 1.5%. In both wet and dry insulin, there is one molecule per unit cell whose space group is $R3$. This suggests a molecule having threefold symmetry but this symmetry need not extend to atomic dimensions in order to agree with the X-ray data. There is very good agreement between the X-ray and ultracentrifuge molecular weights.

A total of fifty-nine reflections were observed, none corresponding to a spacing less than 7.05 Å. This minimum spacing at once imposes a limit on what information can be obtained about the small-scale

* A complete bibliography can be found in Pauling and Niemann's paper [34].
[13] Astbury, W. T., Dickinson, S., & Bailey, K. Biochem. Jour. 29: 2351. 1935.
[14] Pauling, L., & Niemann, C. Jour. Am. Chem. Soc. 61: 1860. 1939.
[15] Bernal, J. D., Fankuchen, I., & Riley, D. Nature 143: 897. 1939.
[16] Riley, D., & Fankuchen, I. Nature 143: 648. 1939.
[17] Langmuir, I. Proc. Phys. Soc. 51: 592. 1939.
[18] Wrinch, D. M. Proc. Roy. Soc. London A161: 505. 1937.
[19] Wrinch, D. M. Trans. Faraday. Soc. 33: 1368. 1937.
[40] Langmuir, I., & Wrinch, D. M. Proc. Phys. Soc. 51: 613. 1939.
[41] McFarlane, A. S., & Kekwick, R. A. Biochem. Jour. 32: 1607. 1938.
[42] Bernal, J. D. Proc. Roy. Soc. London A170: 75. 1939
[43] Wrinch, D. M., & Langmuir, I. Jour. Am. Chem. Soc. 60: 2247. 1938.
[44] Kunitz, M., & Northrop, J. H. Jour. Gen. Physiol. 18: 433. 1935.
[45] Perutz, M. F. Ph.D. Thesis, University of Cambridge. 1939.
[46] Crowfoot, D. & Riley, D. Unpublished data. 1940.

structure of dry insulin. The translation of the intensities into a set of F^2 values is a direct and simple process. Using these, one could compute a three-dimensional Patterson summation, although the task would have been very tedious. Instead, Crowfoot used the data to compute one two-dimensional Patterson projection and four Harker two-dimensional sections. These were computed, referred to an hexagonal unit cell, $a = 74.8$ Å, $c = 30.9$ Å, which contains three identical molecules at the points $(0, 0, 0)$ $(1/3, 2/3, 1/3)$ $(2/3, 1/3, 2/3)$. Due to the rhombohedral symmetry of the crystal these five diagrams are sufficient to map the significant details of the three-dimensional Patterson structure. The horizontal Harker sections are 5 Å apart, sufficiently close together when used in conjunction with the vertical sections to assure that no important peak will be missed. Crowfoot gave in a table (see TABLE 2) the approximate locations of four representative peaks which, when operated upon by the symmetry elements present reproduced the entire structure of Patterson peaks. She was unable to suggest a structure which could explain this peak system. Attempts have been made to explain the Patterson diagrams[42, 43] but due to the rather meagre data upon which any such suggestions must at present be based, they should be considered as tentative.

TABLE 2

APPROXIMATE CO-ORDINATES OF REPRESENTATIVE PEAKS IN THE PATTERSON
DIAGRAM FOR DRY INSULIN (CROWFOOT[13])

	x	y	z
A	.147	.027	.12
	.127	.113	−.25
B	.313	.047	−.17
	.30	.047	.33
C	.29	.19	0
D	0	0	.5

Recently, a preliminary note has been published by Crowfoot and Riley[14] on wet insulin. The internal regularity of these crystals is apparently very good, as the diffraction patterns are very much better than for dry insulin, spacings as low as 2.4 Å being observed. It seems likely that a comparison of the wet and dry insulin projections and sections will permit the identification of the peaks with the respective origins they go with. Such a sorting out of the peaks would be a big step forward towards finding the phases of the F terms. When these are all known, we will know what the X-rays can tell us about the structure of insulin.

To sum up: X-ray studies permit the determination of molecular weights which are in good agreement with those obtained by other methods. An abundance of X-ray data can be obtained to test proposed theories. To be sure, the test at present is a negative one. No theory can be accepted which does not agree with the X-ray evidence, but at present the converse of this statement is, unfortunately, not true. Progress is, however, being made and no doubt X-ray studies will continue to make substantial contributions to the steadily increasing store of knowledge that we have about proteins.

Insulin is the only protein crystal on which any intensive X-ray studies have been published. It would appear desirable that other protein crystals be more extensively studied in order to avoid the danger of using conclusions drawn from studies of a very special hormone, insulin, to symbolize the structure of proteins in general. It is known that hemoglobin[45] and lactoglobulin[46] are being further investigated. The writer hopes to make an X-ray survey of those protein crystals which have not been studied, in order to acquire enough data to permit of the formulation of a classification based upon X-ray measurements, and also in order to determine those proteins which are likely to repay more intensive study.*

* During 1939–1940, preliminary X-ray studies have been made by the writer (as National Research Fellow in Protein Chemistry) on horse serum albumin, ribonuclease (from Dr. M. Kunitz) and urease, as part of this suggested X-ray survey. The results will be published shortly, it is hoped.

THE PRIMARY PROCESS IN PHOTOCHEMISTRY*

By

W. Albert Noyes, Jr., E. W. R. Steacie, Hugh S. Taylor,
Edward Teller, and W. West

CONTENTS

INTRODUCTION TO THE CONFERENCE ON THE PRIMARY PROCESS IN PHOTOCHEMISTRY

By W. Albert Noyes, Jr.

From the University of Rochester, Rochester, New York

Photochemists owe a real debt of gratitude to the officers of the New York Academy of Sciences for granting us the opportunity to get together and discuss our problems. In a field which is developing as rapidly as photochemistry it is quite important that there be a free exchange of ideas. One has only to reflect on the status of this field twenty years ago to realize what strides have been made, not alone in experimental technique, but in the interpretation of the data.

In dealing with polyatomic molecules there is the ever present problem of determining the nature of the primary process. The situation is much more complex than that for diatomic molecules where in many instances the appearance of the spectrum permits a decision to be made between dissociation, predissociation, and activation. Indeed, the detailed analysis of the absorption spectrum of a polyatomic molecule is no easy task, and one of the important problems confronting us lies just there. Frequently it is impossible to tell whether an absorption spectrum is truly continuous or merely appears to be so because of the superabundance of absorption bands. The study of the fluorescence of complex molecules has remained in most instances in a purely qualitative state, partly because of experimental difficulties and partly it has not been realized sufficiently just how valuable the information to be gleaned from such a study really is.

Relatively little is known as yet concerning the transfer of vibrational and electronic energy between polyatomic molecules by collision. Thus fluorescence may give us information not alone about the primary process, but through studies of its quenching about the transfers of energy. In some instances the fluorescence also tells us a great deal about the reaction products.

When we come to secondary processes based on reactions of free radicals and atoms, several points may be mentioned which have been ignored all too frequently. One of these is diffusion to the walls, a process which may compete with homogeneous gas phase reactions providing the latter have activation energies greater than about 10,000 calories. Radical recombination reactions can be studied

through a variation of absorbed light intensity, but it must be kept in mind that unless the incident intensity and the pressure are both varied, two simultaneous variations of variables are being introduced. With intermittent light sources such as sparks and with non-homogeneous beams of radiation it is difficult to be sure of the mechanism of any reaction involving radicals.

Many of these problems cannot be solved satisfactorily in the laboratory, thus leaving to very unsatisfactory indirect methods the determination of the various steps in the mechanism. These indirect methods are based all too often on a species of intuition, and one must admit that rigorous proof of the many things we believe about reaction steps is lacking.

It is unfortunate that even many of the simple photochemical reactions of polyatomic molecules are in this unsatisfactory state, but there is hope for the future in that more care is being exercised in the drawing of conclusions, and a clearer idea exists now concerning the plausibility of various intermediate steps than was possible in the past. It is hoped that the discussions of this symposium will bring to light many instances of carefully planned, quantitative investigations of photochemical reactions.

ASYMMETRIC VIBRATIONS EXCITED BY AN ELECTRONIC TRANSITION

By Edward Teller

From George Washington University, Washington, D. C.

INTRODUCTION

If the equilibrium configuration of a polyatomic molecule has the same symmetry in the fundamental and an excited electronic state, then according to the simplest form of the Franck-Condon principle no asymmetric vibrations will be excited during the electronic transition.[1] Indeed, the forces due to the change in equilibrium configuration are symmetrical and cause only symmetrical vibrations. This fact simplifies the vibrational structure of absorption spectra of polyatomic molecules. The question of asymmetric vibrations has also a direct bearing on problems of photochemistry. To dissociate one atom from a symmetrical molecule a non-symmetrical motion must often take place. Thus if a water, methane, or benzene molecule loses a hydrogen atom, the original symmetry is destroyed. But asymmetric vibrations may become important even if the dissociation corresponds to a symmetrical motion of the molecule. Consider the breaking of the C-C bond in ethane following light absorption. Part of the light energy is transformed into vibrations. Dissociation occurs if enough of the vibration energy happens to be accumulated in the C-C bond.[2] This will happen after a relatively short time if the available energy is distributed between the three symmetrical vibrations of ethane. But it will take much longer if all the eighteen internal degrees of freedom are involved.

INFLUENCE OF ASYMMETRIC VIBRATIONS ON ABSORPTION SPECTRA

A more detailed application of the Franck-Condon principle to spectra of polyatomic molecules shows that asymmetrical vibrations may occur for three reasons. First, the difference of frequency of an asymmetric vibration in the upper and lower electronic states causes a difference in the wave functions of corresponding quantum numbers. Such a frequency difference will, therefore, make possible

[1] Herzberg, G., & Teller, E.   Zeit. Phys. Chem. B21: 410–46.   1933.
[2] Polanyi, M., & Wigner, E.   Zeit. Phys. Chem. A139: 439–52.   1928.

transitions between states of different vibrational quantum numbers. Secondly, an electronic transition may be forbidden if the nuclei are in a symmetrical configuration and may become permitted when an asymmetric vibration destroys the symmetry of the molecule. Then the transition is possible only if the quantum number of at least one appropriate asymmetric vibration changes by at least one quantum. Lastly, the molecules may have in the excited state a different, e. g. a lower, symmetry than in the fundamental state. Then vibrations toward the new equilibrium configurations are excited and these vibrations are asymmetrical with regard to the symmetry of the fundamental state.

The effect of a change in vibrational frequency is important if the fractional change is great and if transitions between higher vibrational levels are involved. In fact the differences of amplitudes of corresponding vibrational levels of the two electronic states increase with increasing quantum number while the differences between amplitudes of consecutive levels in the same electronic state decrease. If finally the amplitude of the v-th vibrational level in the lower electronic state becomes as great as that of the $(v + 1)$-th level in the excited state, the $v \to v$ vibrational transition will cease to have a high intensity compared to transitions in which the quantum number changes. A convenient qualitative description of the relative intensities can be given by calculating the ratio of the intensity of the $v \to v$ transition I_{vv} and the sum of all the intensities $\sum_{v^1} I_{vv^1}$ of all transitions starting from the v-th state. Introducing the quantity ξ which is the geometrical mean of the frequencies in the two electronic states divided by their arithmetical mean, one obtains

$$\frac{I_{00}}{\sum\limits_{v^1} I_{0v^1}} = \xi, \quad \frac{I_{11}}{\sum\limits_{v^1} I_{1v^1}} = \xi^3, \quad \frac{I_{22}}{\sum\limits_{v^1} I_{2v^1}} = \xi\left(\frac{3\xi^2 - 1}{2}\right)^2, \quad \frac{I_{33}}{\sum\limits_{v^1} I_{3v^1}} = $$
$$\xi^3\left(\frac{5\xi^2 - 3}{2}\right)^2.$$

If for instance the frequencies in the two combining electronic states have the ratio 3 : 2, one finds $\xi = 0.98$ and

$$\frac{I_{00}}{\sum\limits_{v^1} I_{0v^1}} = 0.98, \quad \frac{I_{11}}{\sum\limits_{v^1} I_{1v^1}} = 0.94, \quad \frac{I_{22}}{\sum\limits_{v^1} I_{2v^1}} = 0.866, \quad \frac{I_{33}}{\sum\limits_{v^1} I_{3v^1}} = 0.778.$$

Excitation of asymmetrical vibrations due to a change in frequency have been found in the spectra of mercury-halides.[3] These are sym-

[3] Wehrli, M. Helvetica Physica Acta 11: 339. 1938; 13: 153. 1940.

metrical linear triatomic molecules. The spectra consist of a pro-
gression of strong bands which correspond to transitions in the sym-
metrical vibration. In addition weaker bands have been found
which have been attributed to $0 \rightarrow 2$ and $2 \rightarrow 0$ transitions of the
bending vibration (the $0 \rightarrow 1$ and $1 \rightarrow 0$ transitions of this vibration
are forbidden). The bending frequency in the fundamental state
is low: 70 cm^{-1}, 40 cm^{-1}, and 30 cm^{-1} for $HgCl_2$, $HgBr_2$, and HgI_2
respectively. Thus a number of levels are excited at room tempera-
ture and for the higher levels the probability for a change of quantum
number becomes considerable. This explains a few peculiarities of
the spectra. The bands of $HgCl_2$ are fairly broad, those of $HgBr_2$
are broader, and those of HgI_2 are even more diffuse. Actually what
appears on the plates as one band is a superposition of a number of
vibrational transitions. For instance what is labelled as a $0 \rightarrow 0$
transition of the bending frequency consists of the $0 \rightarrow 0$, $1 \rightarrow 1$,
$2 \rightarrow 2$ . . . etc. transitions. These transitions should form a sequence
of bands extending towards the red and their separation would be
the difference between the bending frequencies in the lower and
upper electronic states. Owing to the smallness of this separation
(< 5 cm^{-1}) the bands are not resolved. The Boltzmann factor tends
to make the $0 \rightarrow 0$ band strongest and accordingly the unresolved
band-progression should have a maximum at its violet end. This
seems to be the case in $HgCl_2$, but not in $HgBr_2$ and HgI_2. Actually
the population of the first ($v = 1$) level is greater than that of the
zero$^{\text{th}}$ ($v = 0$) level even in $HgCl_2$, since the vibration is degenerate
and the statistical weight of the v-th level is $v + 1$. The maximum
of the band sequence does not coincide, therefore, with its violet edge
and lies farther toward the red. This effect is particularly strong in
$HgBr_2$ and HgI_2 in which molecules the Boltzmann factor has a
smaller influence.

Similar statements hold for the bands which have been interpreted
as $0 \rightarrow 2$ and $2 \rightarrow 0$ transitions but which, in reality, are band
sequences $0 \rightarrow 2$, $1 \rightarrow 3$ . . . and $2 \rightarrow 0$, $3 \rightarrow 1$, . . . But in these
sequences the intensity of bands increases toward the red for an
additional reason,—the greater probability of exciting asymmetric
vibrations in the higher quantum states. Thus the maximum of these
band sequences will be shifted towards the red by a larger amount
than the maximum of the $0 \rightarrow 0$, $1 \rightarrow 1$ . . . sequence. Now the dis-
tance of the maximum of the $0 \rightarrow 0$, $1 \rightarrow 1$ . . . and $0 \rightarrow 2$, $1 \rightarrow 3$ . . .
sequences has been interpreted in the literature as the separation of
the $0 \rightarrow 0$ and $0 \rightarrow 2$ bands and, therefore, as twice the bending fre-

quency in the excited state. We see that the actual separation of the $0 \rightarrow 0$ and $0 \rightarrow 2$ bands is smaller than the measured separation of the maxima. In a similar manner the separation of the $2 \rightarrow 0$ and $0 \rightarrow 0$ bands is smaller than the separation of the corresponding maxima. Thus the real bending frequency in the upper electronic state is greater than the value given in the literature while the frequency in the fundamental state is smaller. The changes here proposed make the interpretation of the mercury halide spectra more plausible, since according to the uncorrected measurements the bending frequency in $HgBr_2$ and HgI_2 would be twice as small in the excited state as in the fundamental state; such a great ratio corresponding to a factor of four in the force constant is hardly to be expected.

The second reason for the appearance of asymmetric vibrations,— that the transition probability may depend on an asymmetric displacement and may even be different from zero only on account of such a displacement,—is illustrated by the weak absorption system of benzene around $40,000$ cm^{-1}. The transition can be described according to the correspondence principle as a vibration of the charge distribution between two distributions which represent the two Kékulé formulae. It is seen easily that this motion of the charge does not cause a change in the dipole moment and the transition is, therefore, forbidden. A change of dipole moment appears, however, during the motion of the electrons if the regular carbon hexagon is replaced by an elongated hexagon. This distortion corresponds to a vibration of the type $E_g{}^+$. Detailed discussion[4] of the symmetry of benzene shows that this and only this type of vibration can make the transition in question allowed. Experimentally the conclusions are confirmed,—the main bands of the system can be interpreted as $0 \rightarrow 1$ and $1 \rightarrow 0$ transitions of an $E_g{}^+$ carbon vibration on which progressions of the symmetrical carbon vibration and $v \rightarrow v$ transitions of further asymmetric vibrations are superposed.

The last and most effective reason for the excitation of asymmetric vibrations is a change of symmetry of the equilibrium configuration. Such change of symmetry is in general to be expected if the electronic proper function has an orbital degeneracy for the symmetrical nuclear configuration.[5] Spectroscopic evidence for such behavior is difficult to interpret because of the intricate coupling between electronic motion and vibration which must be always expected if electronic

[4] Sponer, H., Nordheim, G., Sklar, A. L., & Teller, E.   Jour. Chem. Phys. **7**: 207.   1939.
[5] Jahn, H. A., & Teller, E.   Proc. Roy. Soc. London **A161**: 220.   1937.

degeneracy depends on the symmetry of the nuclear configuration and if, therefore, the degenerate state is split by asymmetric vibrations.[6] It is probable that the strong absorption system around 55000 cm^{-1} in benzene cannot be interpreted without assuming a nuclear configuration in the excited state whose symmetry is different from that in the ground state.[7] Other reasons also suggest that the upper state of this system is degenerate.[8]

It is interesting to notice in this connection that asymmetric vibrations must be excited in all electronic transitions of molecules with cubic symmetry,—such as CCl_4 or SF_6. Transitions in such molecules are forbidden unless at least one of the two combining states is three-fold degenerate. Thus a transition is either forbidden and then excitation of an asymmetric vibration is needed to make the transition possible, or the transition is allowed, then at least one of the combining levels is degenerate and the symmetrical configuration of the nuclei does not correspond to equilibrium.

The equilibrium configuration may have, of course, a different symmetry in the two combining electronic states even if no degeneracy occurs in the symmetrical configuration. In this case there will be an unstable equilibrium at the symmetrical configuration.

SPLITTING OF DEGENERATE ELECTRONIC STATES BY ASYMMETRIC VIBRATIONS

We shall now investigate in greater detail in what way a degenerate electronic state behaves. Which of the vibrations may cause splitting has been investigated by group theoretical methods[*], but since most of the vibrations that cause the splitting are themselves degenerate it remains to be decided in which direction of the vibration the splitting is strongest and in which direction the slope of the potential surface is steepest. One may then conclude that motion takes place with greatest probability in the direction of steepest slope. The treatment will be carried out only for small displacements of the nuclei from the symmetrical configuration; only terms depending on the homogeneous first power of the displacements will be taken into account. The results can be summarized by stating that for twofold degenerate electronic states the slope of the energy surfaces is the same for all directions of the degenerate vibration which causes the

[6] Renner, R. Zeit. Physik 92: 172. 1934.
[7] Nordheim, G., Sponer, H., & Teller, E., in print.
[8] Goeppert-Mayer, M., & Sklar, A. L. Jour. Chem. Physics 6: 645-52. 1938.
[*] See reference 5.

splitting.   For threefold degenerate electronic states which occur only
in molecules of cubic symmetry* the slope is steepest in twofold degen-
erate vibrations if the displacement preserves a fourfold axis (or in
a tetrahedron a fourfold rotation reflection axis) ; for a threefold
degenerate vibration the steepest slope is obtained for such displace-
ments which preserve a threefold axis.   Of course, the steepest pos-
sible slope will in general occur for displacements involving more than
one normal vibration.   The displacement will often involve vibrations
of more than one symmetry type.   With what coefficients these vibra-
tions are represented in the displacement along the steepest slope
cannot be decided by symmetry arguments.

To prove the above statements we first consider the sum of the
energies of the states into which a degenerate level splits by an
asymmetric displacement of the nuclei.   This sum does not depend
linearly on asymmetric displacements of nuclei, for the same reason
for which the energy of a nondegenerate level does not depend linearly
on such displacements.   In fact, the sum of energies must have the
same value for all displacements which are obtained from each other
by symmetry operations and the sum is, therefore, a totally sym-
metrical representation.   If on the other hand the sum would con-
tain terms depending linearly on a non-totally symmetrical normal
coordinate, the sum would transform according to the same representa-
tion as the normal coordinate.

Secondly, we investigate the sum of squares of energy differences for
an asymmetrical and symmetrical configuration.   We continue to
assume that the energy differences are linear functions of the dis-
placements and the sum of the squares is, therefore, a quadratic func-
tion.   We further consider the dependence of the sum of squares on
one set of degenerate normal coordinates which corresponds to a
representation R of the group of symmetry operations of the molecule.
The sum of squares remains invariant under symmetry operations
and is, therefore, a unit representation of the group.   However, it is
at the same time a homogeneous quadratic function of the degenerate
set of normal coordinates and must, therefore, be one of the repre-
sentations into which the symmetrical direct product $R \times R$ can be
reduced.   Only one unit representation is contained in $R \times R$, namely
the sum of squares of the normal coordinates belonging to the degen-
erate set.   Thus the sum of squares of the energy differences differs
from the sum of squares of the normal coordinates of the degenerate
vibration only in a constant factor.   As a general rule this factor is,
of course, different for different normal vibrations.

---

*We disregard the icosahedron group.

According to the above theorems it is easy to see in what way a twofold degenerate electronic state is split by a degenerate vibration. Since the sum of the energies must remain unchanged, the two energy differences between displaced and undisplaced configurations must be opposite and equal so that for a given direction of displacement the slopes of the two crossing energy surfaces are opposite and equal. Furthermore, the sum of the squares of energy differences is a multiple of the sum of the squares of the normal coordinates. This is possible only if the slopes of the two energy surfaces have the same absolute value for every direction of the degenerate displacement. One
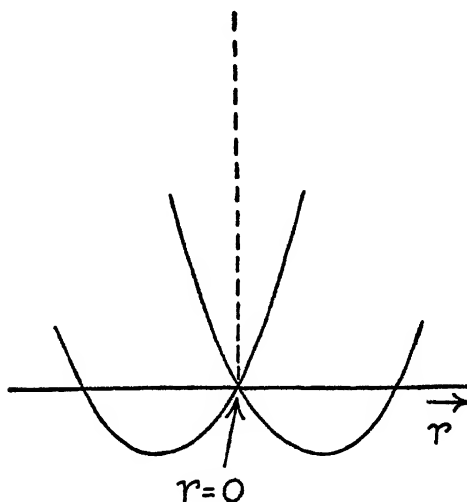
FIGURE 1

can go further in visualizing the energy surface by using the result of the general treatment,* that twofold degenerate electronic states are not split by threefold degenerate vibrations. The splitting is then either due to nondegenerate vibration and one obtains for the energy curves two straight lines of opposite slopes which intersect in the symmetrical configuration, or the vibration is twofold degenerate and then the dependence of the energy surfaces on the two normal coordinates will be given by a circular cone with its apex at the symmetrical configuration. For greater displacement of the nuclei quadratic terms become important and if the molecule is stable in the excited state one may get such energy curves as shown in FIGURE 1.

---

* See 5.

This figure is valid for a nondegenerate vibration. The energy surface for a degenerate vibration may be obtained by rotating FIGURE 1 around the dotted line. This surface is only an approximate one since higher order terms in the energy will cause the surface to deviate from cylindrical symmetry. The surface with its circular trough around $r = 0$ is a good approximation only if the trough is at small r values. There is reason to believe that this is the case of the upper state of the 55000 cm^{-1} band in benzene.

For threefold degenerate electronic states the statements about the sum of energy differences and the sum of their squares do not determine the energy surface in a unique way. It is easy to see, however, that the sum of the differences being zero and the sum of the squares being a quantity which is independent of the direction, the maximum slope is obtained if two energy differences are equal and the third is twice as great and has the opposite sign. This will happen if the asymmetric displacement preserves symmetry elements which can cause an energy level to be twofold degenerate. This means that at last a threefold or a fourfold axis must be retained after the displacement. Now one can show that in a tetrahedral or octahedral molecule a twofold degenerate vibration always destroys the threefold axis. But the vibration in which the two opposite edges of the tetrahedron become shorter while the other four edges are extended preserves the fourfold reflection rotation axis joining the shortened edges. A maximum slope will be expected for this vibration. In a similar way a twofold degenerate vibration of an octahedral molecule preserves the fourfold axis if the two atoms on that axis approach the center of the octahedron while the four other corners of the octahedron move away from the center. But the relevant threefold degenerate vibration in a tetrahedron or an octahedron never preserves a fourfold axis but may preserve a threefold one. Thus the motion of the central atom in a tetrahedron towards one of the corners preserves trigonal symmetry; and, in an octahedron the threefold axis is preserved if opposite triangles of the octahedron contract in a symmetrical way while the edges joining the corners of the triangles to each other are extended. One can show that the vibrations just described are the ones with appropriate symmetry to split threefold degenerate electronic levels.

In order to obtain the detailed dependence of the energy surfaces on the nuclear displacements we consider a tetrahedral AB_4 molecule. We choose the x, y, and z coordinate axes along the fourfold reflexion rotation axes of the tetrahedron. The twofold degenerate vibration

which preserves the fourfold axis along x will also preserve the two reflexion planes which contain the x axis and lie at $45°$ to the y and z axes. A threefold degenerate electronic level splits under the displacement considered into three levels two of which are symmetrical to one of the two planes just mentioned while the third may be either symmetrical to both planes or anti-symmetrical to both. The first wave functions have the same energies while the third has (-2)-times that energy. We may choose ψ_x, ψ_y and ψ_z as the degenerate wave functions; according to the representation to which the electronic wave functions belongs ψ_x may be symmetrical to both planes through the x-axis, or antisymmetrical to both. ψ_y and ψ_z behave in an analogous manner with regard to the y and z axes. Denoting by η_x the amplitude of the twofold degenerate displacement which preserves the fourfold axis along x, we obtain for the perturbation matrix

$$C\eta_x \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where C is a constant and the three rows and columns of the matrix refer to ψ_x, ψ_y and ψ_z so that $C\eta_x$ is the perturbation energy for ψ_y and ψ_z and $2C\eta_x$ is the perturbation on ψ_x. The corresponding matrices for the vibrations that preserve the fourfold reflexion—rotation axes along y and along z are

$$C\eta_y \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } C\eta_z \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

respectively. It will be noticed that for $\eta_x = \eta_y = \eta_z$ the sum of the perturbations is zero as should, indeed, be expected since the three displacements on each atom are at $120°$ to each other and add up to zero. Since all three perturbation matrices are in diagonal form their sum gives immediately the perturbation energy on ψ_x, ψ_y and ψ_z as a function of the twofold degenerate displacements. If the displacements are represented by two coordinates $\eta = \eta_x$ and $\eta^1 = \dfrac{1}{\sqrt{3}}$ $(\eta_y - \eta_z)$ in a plane (η, η^1) the three energy surfaces are three planes which are inclined at equal angles to the plane (η, η^1). These three planes pass through the origin $\eta = \eta^1 = 0$ corresponding to the symmetrical configuration. They may be obtained from each other by rotations of $120°$ around an axis through the origin and perpendicular to the plane (η, η^1).

The splitting of a threefold degenerate electronic state (ψ_x, ψ_y, ψ_z) by a threefold degenerate vibration may be studied by using the vibration of the A-atom against the B_4 tetrahedron as an example for the vibration. Calling the displacements of the A atom along the three axes ξ_x, ξ_y and ξ_z one obtains the three perturbation matrices*

$$k\,\xi_x \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad k\,\xi_y \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } k\,\xi_z \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

respectively. Here k is again a constant. That the perturbation is given by these matrices can be verified by considering, for instance, the displacement ξ_x. The proper functions of the corresponding matrix are ψ_x, $\dfrac{1}{\sqrt{2}}(\psi_y + \psi_z)$ and $\dfrac{1}{\sqrt{2}}(\psi_y - \psi_z)$ with the proper energies 0, $k\xi_x$ and $k\xi_x$ respectively. Now ξ_x preserves the same planes of reflection as the twofold degenerate displacement η_x and it can be easily verified that $\dfrac{1}{\sqrt{2}}(\psi_y \pm \psi_z)$ are the wave functions which are symmetrical to one of these two planes, while ψ_x is symmetrical to both (or none) of the planes. One also can see that the energy of ψ_x does not change linearly with ξ_x while the energies of $\dfrac{1}{\sqrt{2}}(\psi_y + \psi_z)$ and $\dfrac{1}{\sqrt{2}}(\psi_y - \psi_z)$ are opposite and equal.

The matrix for an arbitrary displacement with the three components ξ_x, ξ_y, and ξ_z is

$$k \begin{pmatrix} 0 & \xi_z & \xi_y \\ \xi_z & 0 & \xi_x \\ \xi_y & \xi_x & 0 \end{pmatrix}$$

The proper value problem for this matrix leads to an equation of the third degree the roots of which give the energy change ΔE as a function of the components ξ_x, ξ_y, and ξ_z

$$\Delta E = k\left(\xi_x\xi_y\xi_z + \sqrt{\xi_x^2\xi_y^2\xi_z^2 - \frac{1}{27}(\xi_x^2 + \xi_y^2 + \xi_z^2)^3}\right)^{\frac{1}{3}} + \text{ conj. compl.}$$

One can verify with the help of this expression that the maximum ΔE for a given value of the displacement $(\xi_x^2 + \xi_y^2 + \xi_z^2)^{\frac{1}{2}} = \xi$ is obtained if $\xi_x = \xi_y = \xi_z = \dfrac{1}{\sqrt{3}}\xi$ *i. e.* if the displacement is along a

---

* The signs of the matrices are correct if one corner of the tetrahedron is in the octant of positive x, y, and z.

trigonal axis. For this case the expression simplifies to

$$\Delta E = \frac{2}{\sqrt{3}} k\xi$$

with two other roots which are given by

$$\Delta E = -\frac{1}{\sqrt{3}} k\xi$$

It is of some interest to notice that breaking of a bond in a tetrahedral molecule involves motion along a trigonal axis.

ASYMMETRIC VIBRATIONS OF GREAT AMPLITUDE

While excitation of asymmetric vibrations by one or two quanta gives rise to characteristic bands in spectra and may lead to the interpretation of a band system, the photochemist is primarily concerned with the cases where asymmetric displacements reach great amplitudes. Dissociation often requires such great amplitudes; also in the distribution of energy between different degrees of freedom the asymmetric vibrations become important only if they acquire several quanta.

From the reasons discussed earlier, excitation of an asymmetric vibration in a forbidden electronic transition changes the quantum number of the asymmetric vibration only by one unit and excitation of this type is, therefore, of no importance in our present discussion.

The most effective way of exciting asymmetric vibrations is by an electronic transition in which the final state has a lower symmetry than the initial state. As a general rule this is the case if the final state has an orbital degeneracy. But then only certain asymmetric vibrations will be excited, namely, those which split the degeneracy. As an example, we consider the C_2H_6 molecule. An absorption with its transition moment perpendicular to the C-C axis leads to a degenerate electronic state. This state can be split, however, only by a degenerate vibration. Moreover, if we assume for the hydrogens a staggered equilibrium configuration the vibration which causes the splitting must be symmetrical to the symmetry center of the molecule, and if we assume the hydrogens to be in opposition only those vibrations are effective which are symmetrical to the plane bisecting at right angles the C-C axis. There are thus only three degenerate vibrations which will be strongly excited and they are of such a symmetry that together with the symmetrical vibrations they cannot account for the breaking of one C-H bond.

It is always possible, as has been pointed out, that potential surfaces of other asymmetric vibrations have maxima rather than minima in the excited state. Such vibrations will get strongly excited in a time comparable to their vibrational period in the fundamental state,—that is the time required for the wave function representing the zero point vibration in the fundamental state to flow apart once the confining potential is removed.

But in order to excite an asymmetric vibration strongly it is not necessary to convert its potential minimum into a maximum. It is sufficient to make the minimum more, or less, shallow. As a general rule the accompanying change of frequency causes in band spectra only changes of the vibrational quantum number by two units. But a strong excitation of the symmetrical vibrations may cause an enhanced excitation of the asymmetrical vibrations through a strong anharmonic coupling. That can be seen in the following semiclassical way. During the strong symmetrical vibrations the asymmetric frequencies may change considerably and this will give rise to transfer of energy to the asymmetric vibrations. The transfer will be considerable if during one vibrational period of the asymmetric vibration its frequency is greatly changed by the displacements which the symmetrical vibrations have suffered in the meantime. This condition is satisfied quite often; one may see, for instance, that the bending vibrations of the CH_3 groups in C_2H_6 will be quite strongly influenced if the two CH_3 groups vibrate against each other with a great amplitude.

A more exact quantum theoretical description of the excitation of asymmetric vibrations through strong symmetrical vibrations can be given with the help of accidental,—or Fermi-Dennison type,—degeneracies. If a sum of frequencies $\nu_1 + \nu_2$ differs by a small amount from another frequency ν_3, proper functions are obtained that have partly the properties of a one quantum state of both ν_1 and ν_2 and partly the properties of a one quantum state of ν_3. Similar relations hold if a sum of several frequencies (or multiple frequencies) agrees with a sum of other frequencies (or multiple frequencies). The intermixing of quantum states becomes weaker if more frequencies are involved but it increases with the anharmonic forces and, therefore, with the amplitudes. As the vibrations grow stronger anharmonic mixing of more and more frequencies becomes possible and practically each vibrational level will become "accidentally" degenerate with other levels. Though symmetry restricts the levels that can perturb each other, this restriction still leaves many levels that

can interact and, roughly speaking, its effect is that asymmetric vibrations can be excited through anharmonic forces only by even quantum numbers.

If in an absorption spectrum many vibrations are excited with great amplitudes, the spectrum may become too complicated to be analyzed. In photochemistry such complicated spectra might be, however, very interesting. In some bands principally symmetrical vibrations may be excited, in others, which may lie quite near to the first, asymmetric vibrations may be strongly involved. This may have the consequence that absorption in different bands may lead to very different dissociation products or even that absorption in some of the bands leads to dissociation while absorption in others does not. All this may happen in a single band system belonging to one definite electronic excitation. But details of the molecular vibrations can have such selective effects only if light absorption takes place at relatively low pressure so that the molecule will dissociate or else radiate its energy before colliding with other molecules. In collisions between molecules the specific differences between symmetrical and asymmetrical vibrations disappear. At the same time any vibration of great amplitude will soon distribute its energy among different molecules.

DISCUSSION

The following discussion pertaining to material in this paper was contributed by members of the conference whose names and remarks appear below.

M. BURTON:[*]—According to a calculation by Kimball, ethane formed by bimolecular combination of C_2H_6 has a half-life of $\sim 10^{-12}$ sec. In this calculation Kimball assumed symmetrical approach of the two CH_3 radicals. Apart from this severe limitation introduced in Kimball's calculation, the paper by Professor Teller seems to suggest (by the inverse of the method) that there is a ready transfer of energy between the C-C and other bonds and that consequently the reaction $CH_3 + CH_3 \rightarrow C_2H_6$ might be expected to occur readily. The difficulty in this case is wholly experimental. The mirror experiments of Paneth, Hofeditz, and Wunsch at elevated temperatures show clearly that whatever combination of methyl radicals does occur at pressures less than 10 millimeters is wholly on the wall. Furthermore, it has been found repeatedly that in those cases where such a combination has been suggested (*cf.* the chain-ending step in acetalde-

[*] New York University, New York City, New York.

hyde decomposition) the evidence is better interpreted in terms of another mechanism.

On the other hand, there is little question about the relatively long life of $HgCH_3$, for the reaction between Hg and CH_3 has been repeatedly demonstrated to proceed in the vapor phase.

It would seem consequently that if this interpretation of the chemical evidence is correct, the theoretical considerations are perhaps oversimplified. It may be, for example in the case of C_2H_6, that there are restrictions on the transfer of energy between quantum states of the various degrees of freedom which actually prevent internal stabilization of that molecule when formed by collision of two methyl radicals.

EDWARD TELLER:—The $N_2O + N \rightarrow N_2 + ON_2$ reaction. It would be interesting to study the transfer of an oxygen atom from an N_2O molecule to an N_2 molecule. It is known that N_2O dissociates at high temperatures into N_2 and O. This dissociation is accompanied by a change of multiplicity. Indeed the ground state of N_2O is a singlet and the same holds for N_2 while the lowest state of O is a triplet. By studying the $N_2O + N_2 \rightarrow N_2 + ON_2$ exchange reaction no change of multiplicity takes place. Thus comparing the exchange reaction with the dissociation, information may be obtained about the influence of the change in multiplicity.

The exchange reaction could be studied easily by mixing N_2O with heavy N_2 (or with radiactive N_2) heating the mixture to the reaction temperature, freezing N_2O out and determining the amount of isotopic nitrogen in N_2O.

F. O. RICE:*—Concerning the exchange reaction suggested by Dr. Teller,—

$$N_2O + N_2 = N_2 + N_2O,$$

some other reactions which might be considered in this connection are:

$$CO + CO_2 = CO_2 + CO$$
$$CH_2 = CO + CO = CO + CH_2 = CO$$
$$CH_2 = N \equiv N + N_2 = N_2 + CH_2 = N \equiv N$$
$$CH_2 = N \equiv N + CO = CH_2 = CO + N_2.$$
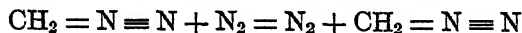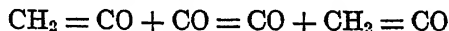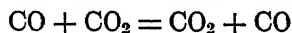
Of these, the last has been actually observed by Staudinger and Kupfer in 1912.

---

* The Catholic University of America.

PHOTOSENSITIZATION EXPERIMENTS
WITH VARIOUS METAL VAPOURS*

By E. W. R. STEACIE

From the National Research Laboratories, Ottawa, Canada

INTRODUCTION

Photosensitized reactions present a much closer analogy to thermal processes than do direct photochemical reactions, since the reactant molecule receives the energy necessary for reaction by collision. Furthermore, the situation is much simpler in a photosensitized reaction than it is in a thermal reaction, since we have a rather definite input of energy to the reactant molecules instead of the Boltzmann distribution.

The present paper will be confined to photosensitization by metal vapours. Of these mercury is by far the simplest from an experimental point of view, and a great deal of work on mercury photosensitization has been done. Mercury possesses two important resonance lines at 1849Å and 2537Å. Almost all the work which has been done, however, has been with λ 2537, which involves a transition from the normal 6^1S_0 state to the 6^3P_1 state. The excitation energy of this state is 4.87 volts, or 112 Kcal. per mole. This is higher than the activation energy of most reactions, and hence a wide variety of processes may be initiated by this means. Some examples are the reactions.

$$H_2 \rightarrow 2H - 103 \text{ Kcal.}[1]$$

$$C_2H_6 \rightarrow C_2H_5 + H - ca. \ 95 \text{ Kcal.}[2]$$
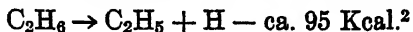
As may be seen from the heats of the above reactions, there is a quite efficient utilization of the excitation energy of the mercury atom, since reactions requiring a hundred kilocalories or more may be initiated. Actually, however, there is some doubt concerning the exact amount of energy available for reactions of the above type. It appeared to be definitely established that at least a large part of the dissociation of hydrogen occurred not by the direct reaction

$$Hg \ (^3P_1) + H_2 \rightarrow Hg \ (^1S_0) + 2H,$$

but in steps,[3]

$$Hg\ (^3P_1) + H_2 \rightarrow HgH + H$$

$$HgH \rightarrow Hg\ (^1S_0) + H.$$

It has recently been found by Olsen,[4] however, that it is not possible to produce resonance excitation of the HgH bands. He concludes that HgH is not present in the normal state, but is formed in excited states by secondary processes involving hydrogen atoms. It is probable, therefore, that the net reaction is

$$H_2 \rightarrow 2H - 103\ Kcal.$$

rather than

$$Hg + H_2 \rightarrow HgH + H - 94.5\ Kcal.$$

$$HgH \rightarrow Hg + H - 8.5\ Kcal.$$
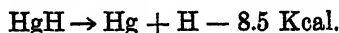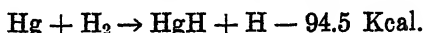
There is also some evidence that we may have similarly

$$C_2H_6 + Hg\ (^3P_1) \rightarrow HgH + C_2H_5.$$
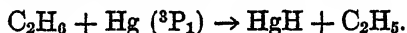
If this is the case, there will be available by mercury photosensitization sufficient energy to split off hydrogen from a molecule provided the overall process is not endothermic to an extent greater than $112 + 8.5 = 120.5$ Kcal.

CADMIUM AND ZINC

In view of the interest which has attached to investigations with mercury, it seemed desirable to extend the work to other sensitizers, and thus to vary the energy input to the reactant molecule. From an experimental point of view only a few substances have excitation energies of the right order of magnitude, together with sufficiently high vapor pressures at reasonably low temperatures. Cadmium and zinc appear to be the most promising of these, and we have commenced a series of investigations of photosensitization by cadmium and zinc.

The energy levels, etc., of cadmium and zinc are, of course, very similar to those of mercury. There is, however, one important difference. As we have seen, Olsen's work makes it very probable that with hydrogen the primary step in the presence of excited mercury atoms is

$$Hg\ (^3P_1) + H_2 \rightarrow Hg\ (^1S_0) + 2H$$

---
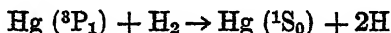
[3] Gaviola, E., & Wood, R. W. Phil. Mag. 7. series 6:1191. 1928. Beutler, H., & Rabinowitch, E. Zeit. Phys. Chem. 88: 231. 1930. Rieke, F. F. Jour. Chem. Phys. 4: 513. 1936.

[4] Olsen, L. O. Jour. Chem. Phys. 6: 307. 1938.

rather than

$$Hg\ (^3P_1) + H_2 \rightarrow HgH + H$$
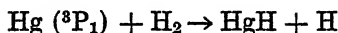
In the case of cadmium,—hydrogen mixtures, however, Bender[5] has observed strong resonance excitation of the CdH bands. This indicates that CdH is formed in the normal state, and suggests that the primary step is

$$Cd\ (^3P_1) + H_2 \rightarrow CdH + H$$

It is therefore not at all certain *a priori* whether with reactions involving a C-H bond split in a hydrocarbon molecule we have in a given case

$$M\ (^3P_1) + HX \rightarrow M\ (^1S_0) + H + X$$

or

$$M\ (^3P_1) + HX \rightarrow MH + X$$
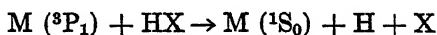
There is therefore some doubt about the amount of energy required to split the C-H bond in a given case. The energies associated with the excited atoms, etc., are summarized in TABLE 1.

TABLE 1

EXCITATION ENERGIES, ETC.

Substance	Resonance Line, Å	Transition	Energy of excited atom, Kcal.	Heat of dissociation of hydride, Kcal.	"Maximum energy available" to split off H atom, assuming intermediate hydride formation, Kcal.
Mercury	1849	$6^1S_0–6^1P_1$	153.9	8.5	162.4
	2537	$6^1S_0–6^3P_1$	112.2	8 5	120.7
Cadmium	2288	$5^1S_0–5^1P_1$	124.4	15.5	139 9
	3261	$5^1S_0–5^3P_1$	87.3	15 5	102.8
Zinc	2139	$4^1S_0–4^1P_1$	133.4	23.1	156.5
	3076	$4^1S_0–4^3P_1$	92.5	23.1	115.6

Experimental

The only previous investigation on cadmium or zinc photosensitizaiton was a brief qualitative investigation with cadmium by Bates and Taylor.[6] This work was not followed up mainly because of the lack of suitable sources of resonance radiation of reasonably high intensity.

[5] Bender, P. Phys. Rev. **36**: 1535. 1930.
[6] Bates, J. R., & Taylor, H. S. Jour. Am. Chem. Soc. **50**: 771. 1928.

We have recently developed lamps[7] similar to high voltage mercury lamps, and these have proved very satisfactory. Spectrograms of a cadmium resonance lamp are shown in FIGURE 1.

Results

ETHANE

We have already reported the results of an investigation of the cadmium photosensitized reactions of ethane.[8] It was found that hydrogen was formed in large quantities in the early stages of the reaction. It thus appears certain that the primary reaction involves a C-H bond split, as is the case in the mercury photosensitized reaction. There are two ways in which this might occur, *viz.*

(1) $C_2H_6 + Cd\ (^3P_1) \rightarrow C_2H_5 + H + Cd\ (^1S_0)$

(2) $C_2H_6 + Cd\ (^3P_1) \rightarrow C_2H_5 + CdH.$
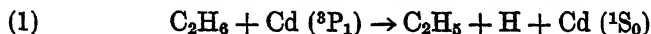
Since the formation of cadmium hydride has been shown to occur in the quenching of $Cd\ (^3P_1)$ resonance radiation by hydrogen, it is obviously possible for it to be formed here. If the reaction occurs by (1), the energy available to split the C-H bond is the excitation energy of the cadmium atom, 87.3 Kcal., plus perhaps a certain amount of kinetic energy, etc. Estimates of the C-H bond strength vary greatly[9, 10] but it seems very unlikely that it is as low as this. Since the heat of formation of CdH from the atoms is 15.5 Kcal., reaction (2) is possible provided that the strength of the C-H bond in ethane is not much over 102.8 Kcal. The primary step thus appears to be reaction (2), and is thus analogous to that which occurs in the mercury photosensitized reaction of ethane.[2] To confirm this point a search is being made for the CdH bands in emission in cadmium-ethane mixtures illuminated by a cadmium-hydrogen discharge.

The subsequent secondary reactions are also similar to those occurring with mercury photosensitization, leading qualitatively to the same products, hydrogen, methane, propane, and butane. The quantum yield is also of the same order of magnitude (ca. 0.3).

It may be noted that this work furnishes an upper limit for the

[7] Steacie, E. W. R., & Potvin, R. Can. Jour. Research B16: 337. 1938. Steacie & Habeeb, unpublished.

[8] Steacie, E. W. R., & Potvin, R. Jour. Chem. Phys. 7: 782. 1939.

[9] Rollefson, G. K., & Burton, M. "Photochemistry and the Mechanism of Chemical Reactions." Prentice-Hall, Inc., New York. 1939.

[10] Burton, M. Jour. Chem. Phys. 6: 818. 1938; 7: 1072. 1939.
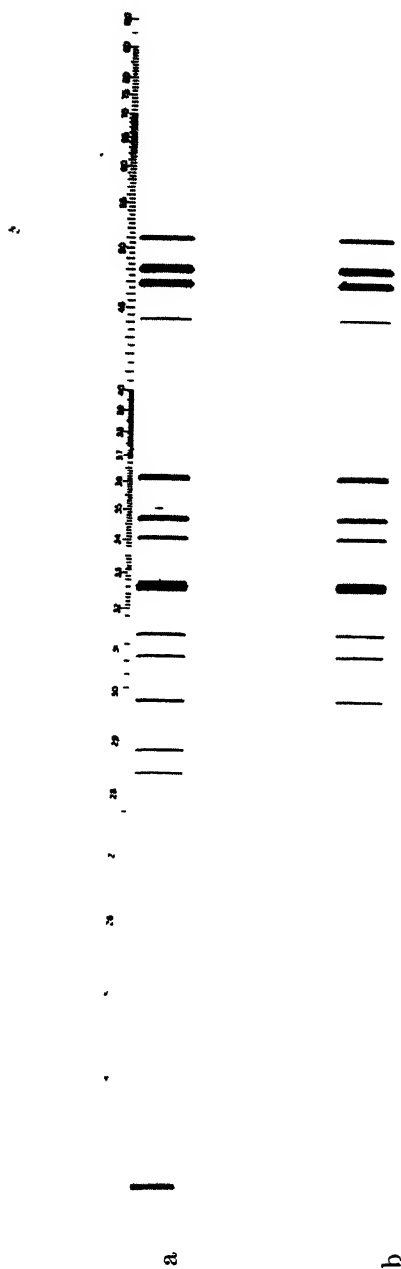
FIGURE 1 (a) Cadmium resonance lamp (of quartz) giving λ 3261 and λ 2288 (b) The same lamp with a Corex D filter. The lower resonance line has been completely removed. It should be noted that the intensity of λ 2288 is higher than appears from the photograph, since the plate sensitivity has fallen off strongly at 2288Å.

strength of the C-H bond in ethane of (102.8 Kcal. + kinetic and internal energy). From specific heat data, etc., it appears that this limit is not over 105–106 Kcal.
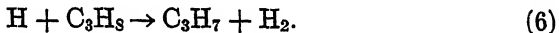
PROPANE

Investigations have recently been made of the mercury[11] and cadmium[12] photosensitized reactions of propane.

There appears to be no doubt that in both cases the primary step is a C-H bond split, *i. e.*

$$C_3H_8 + X \ (^3P_1) \rightarrow C_3H_7 + XH \ (\text{or } X + H). \tag{3}$$
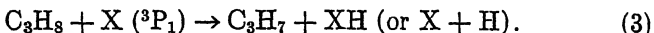
There is, however, one distinct difference between the results of the two investigations. With mercury photosensitization the products are virtually exclusively hydrogen and hexanes, even at temperatures up to 325° C. This would seem to indicate that the only important secondary processes are

$$2 \ C_3H_7 \rightarrow C_6H_{14} \tag{4}$$

$$2 \ H \rightarrow H_2 \tag{5}$$

$$H + C_3H_8 \rightarrow C_3H_7 + H_2. \tag{6}$$

The reaction of hydrogen atoms, produced by a discharge tube, with propane have been investigated by Steacie and Parlee.[13]
They found methane to be the main product at low temperatures, presumably due to the occurrence of

$$H + C_3H_7 \rightarrow C_2H_5 + CH_3 \tag{7}$$

$$H + C_2H_5 \rightarrow 2 \ CH_3 \tag{8}$$

At higher temperatures, there were indications that the propyl radical might be decomposing by

$$C_3H_7 \rightarrow C_2H_4 + CH_3. \tag{9}$$

However, in the case of mercury photosensitization the concentration of hydrogen atoms is very low, and that of propane relatively very high, thus making (7) and (8) unimportant compared with (6). Furthermore recombination reactions of radicals will be greatly favored by the higher pressure, thus diminishing the chance of (9) oc-

[11] **Steacie, E. W. R., & Dewar, D. J.**   Jour Chem. Phys. 8: 571.   1940.
[12] **Steacie, E. W. R., LeRoy, D. J., & Potvin, R.**   Jour. Chem Phys. 9: 306.   1941.
[13] **Steacie, E. W. R., & Parlee, N. A. D.**   Trans. Faraday Soc 35: 854.   1939; Can J. Research B17: 371   1939

curring. There is also the possibility that at higher temperatures the formation of hexane may occur to a large extent by the reaction,

$$C_3H_7 + C_3H_8 \rightarrow C_6H_{14} + H. \tag{10}$$
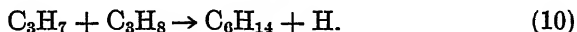
The results of the investigation of the reaction of hydrogen atoms with propane are thus not in any way incompatible with the results of the mercury photosensitized reaction. There is, however, a sharp contrast between these results and those of the cadmium photosensitized reaction. In the cadmium photosensitized reaction at 285° C. hydrogen, methane, butane, and hexanes are formed in large amounts. It therefore appears that there must have been considerable C-C bond splitting, either by reactions (7) and (8), or by reaction (9). Methane could then result from

$$CH_3 + H \rightarrow CH_4 \tag{11}$$

or

$$CH_3 + C_3H_8 \rightarrow CH_4 + C_3H_7 \tag{12}$$

and butane by

$$2\ C_2H_5 \rightarrow C_4H_{10} \tag{13}$$

or

$$CH_3 + C_3H_7 \rightarrow C_4H_{10}. \tag{14}$$
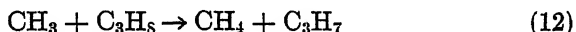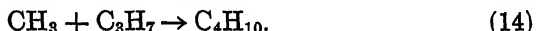
It is, however, difficult to explain the occurrence of reactions (7), (8), and (9) to a large extent in the cadmium photosensitized reaction, when they do not occur in the mercury photosensitized reaction. It is possible that the difference is due to experimental conditions. In the experiments done thus far on cadmium photosensitization, the lamp and reaction vessel have been constructed as a single unit, and it has thus been necessary to use cadmium vapor pressures in the reaction vessel equal to those which will give satisfactory operation of the lamp. This is rather high, *viz.* 0.01 — 0.02 millimeter. Under those circumstances a large fraction of the resonance radiation will be absorbed near the wall, and the wall effects may be responsible for the difference in the products. In view of the comparatively long life of the CdH molecule, the possibility must not be overlooked that the reactions of CdH with hydrocarbon molecules or radicals might occur with relatively different rates compared with those of hydrogen atoms.

BUTANE

Preliminary results on the mercury[11] and cadmium[15] photosensitized reactions of butane indicate a close analogy with those of

[11] **Winkler, & Hay**, private communication.
[15] **Steacie, E. W. R., & LeRoy, D. J.**, unpublished.

propane.   Thus with mercury photosensitization the products are almost entirely hydrogen and octanes, while with cadmium photosensitization a variety of hydrocarbons are formed.

ETHYLENE

The most interesting reaction which we have so far encountered in this work is the photosensitized polymerization of ethylene. A number of investigations of the mercury photosensitized reactions of ethylene have been made, largely by Taylor and his collaborators. Virtually all investigators adopt for the primary step the suggestion of Taylor and Bates[16, 17]

$$C_2H_4 + Hg\ (^3P_1) \rightarrow C_2H_2 + H_2 + Hg\ (^1S_0)$$
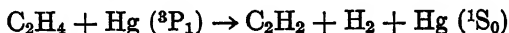
It is then assumed that the subsequent reactions are the photosensitized polymerization of acetylene, together with reactions of hydrogen atoms and of radicals with ethylene, $viz.$

$$H + C_2H_4 \rightarrow C_2H_5$$
$$C_2H_5 + C_2H_4 \rightarrow C_4H_9,\ \text{etc.}$$
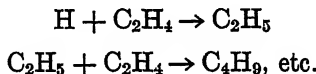
The main evidence for this mechanism is the detection of large amounts of acetylene in the products, and the fact that the pressure $rises$ in the early stages of the reaction and then falls rapidly.

Numerous investigations have also been made of the mercury photosensitized reactions of ethylene-hydrogen mixtures.[17, 18, etc.] The most comprehensive work is that of Jungers and Taylor.[19] They found butane to be the main product together with small amounts of ethane. They suggested that the main reactions occurring are:

$$H + C_2H_4 \rightarrow C_2H_5$$
$$2\ C_2H_5 \rightarrow C_4H_{10}$$
$$2\ C_2H_5 \rightarrow C_2H_4 + C_2H_6$$
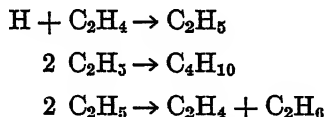
They found approximately the same quantum yield for ethylene alone and for ethylene-hydrogen mixtures.

The first qualitative investigation of the reaction by cadmium photosensitization was made by Bates and Taylor,[20] who reported that in ethylene-hydrogen-cadmium vapor mixtures illuminated with

[16] Taylor, H. S., & Bates, J. R.   Proc Nat Acad Sci 12: 714   1926.
[17] Bates, J. R., & Taylor, H. S. Jour Am Chem Soc 49: 2438   1927.
[18] Taylor, H. S., & Hill, D. G. Jour Amer Chem Soc. 51: 2922. 1929.
[19] Jungers, J. C., & Taylor, H. S. Jour Chem. Phys 6: 325   1938.
[20] Bates, J. R., & Taylor, H. S. Jour Am. Chem. Soc. 50: 771   1928.

λ 3261 no hydrogenation to ethane occurred, but that some polymerization of ethylene took place. We have recently investigated the reaction in some detail.[21] The results are striking in that they indicate a quantum efficiency of about 0.5 for a 1:1 ethylene-hydrogen mixture, but a yield of only about 0.015 for pure ethylene. Furthermore this is not due to lack of quenching by ethylene since qualitative indications are that ethylene quenches strongly.

It is difficult at first sight to reconcile these results with the generally accepted mechanism for the primary step in the mercury photosensitized reaction,

$$C_2H_4 + Hg\ (^3P_1) \rightarrow C_2H_2 + H_2 + Hg\ (^1S_0).$$

It is, of course, possible that ethylene molecules might quench $^3P_1$ cadmium atoms merely to the metastable $^3P_0$ state, and hence receive far too little energy to cause reaction. However, even if this were the case, $^3P_0$ atoms would be expected to cause reaction themselves. Furthermore, considering the relatively long life of the $^3P_0$ state, the small energy difference between the $^3P_0$ and $^3P_1$ states (1.5 Kcal.), and the high temperature, it appears likely that if quenching merely occurred to the the $^3P_0$ level, most of the metastable atoms would be raised to the $^3P_1$ level by collisions of the first kind, and hence the net quenching efficiency would be low. Since this is not the case, it appears probable that ethylene molecules quench the excited cadmium atoms to the ground state, and thus receive a relatively large amount of energy.

Two possible explanations suggest themselves for the difference between the results with mercury and with cadmium.
The reaction

$$C_2H_4 \rightarrow C_2H_2 + H_2$$

is endothermic to the extent of 42.9 Kcal.[22] The activation energy of the reaction has not been determined, but Taylor and van Hook[23] have found that of the reverse reaction to be 42 Kcal. (It has also been roughly estimated at 46.4 Kcal. by Sherman and Eyring.[24] Combining these results we obtain for the activation energy of the ethylene decomposition into acetylene and hydrogen a value of approximately 85 Kcal. It is therefore possible that the 112 Kcal. as-

[21] Steacie, E. W. R., & Potvin, R. Can. Jour. Research B16: 337. 1938; B18: 47. 1940.
[22] Bichowsky, F. R., & Rossini, F. D. "Thermochemistry of Chemical Substances." Reinhold Publishing Corp. New York. 1936.
[23] Taylor, H. A., & van Hook, A. Jour. Phys. Chem. 39: 811. 1935.
[24] Sherman, A., & Eyring, H. Jour. Am. Chem. Soc. 54: 2661. 1932.

sociated with a $^3P_1$ mercury atom can cause reaction, while the 87 Kcal. excitation energy of a $^3P_1$ cadmium atom cannot be used efficiently enough to provide the 85 Kcal. or so which are necessary.

On the other hand there is a great deal of evidence that the polymerization of ethylene can be sensitized by free radicals derived from other sources.[25, 26, 27, 28, 29, 30, 31] Also recent work by Pease and Burnham[32] shows that the thermal polymerization of ethylene can be inhibited by nitric oxide, and thus suggests a free radical mechanism for the thermal reaction. We may therefore consider the possibility that the mercury photosensitized reaction might proceed by a free radical mechanism, with the primary step

$$C_2H_4 + Hg(^3P_1) \rightarrow C_2H_3 + H + Hg(^1S_0).$$

If this were the case, the analogous reaction in the case of cadmium

$$C_2H_4 + Cd(^3P_1) \rightarrow C_2H_3 + H + Cd(^1S_0) \qquad (a)$$

or

$$C_2H_3 + Cd\ H \qquad (b),$$
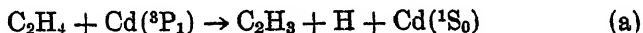
would be ruled out on thermochemical grounds if the C — H bond strength in ethylene were appreciably greater than (a) 87 Kcal, or (b) 103 Kcal. Rice[33] has estimated a value of 104 Kcal, for this bond strength, but the estimate is very uncertain. However, the main argument against this explanation is that the mercury photosensitized polymerization of ethylene proceeds with an initial pressure *increase*, and it seems almost impossible to construct a free radical mechanism which will explain this. Furthermore acetylene is an important product of the reaction. The difference in temperature between the cadmium and mercury investigations does not appear to offer an explanation, since in the only results reported at high temperatures for the mercury photosensitized reaction, Melville,[34] found that there was still an initial pressure rise at 280° C.

The most likely explanation of the results seems therefore to be a

[25] Joris, G., & Jungers, J. C. Bull. Soc. Chim. Belg. 47: 135. 1938.
[26] Rice, O. K., & Sickman, D. V. Jour. Am. Chem. Soc. 57: 1384. 1935.
[27] Sickman, D. V., & Rice, O. K. Jour. Chem. Phys. 4: 608. 1936.
[28] Taylor, H. S. Proc Am. Phil. Soc. 65: 90. 1926.
[29] Taylor, H. S., & Emeleus, H. J. Jour. Am. Chem. Soc. 52: 2150. 1930.
[30] Taylor, H. S., & Jones, W. H. Jour. Am. Chem. Soc. 52: 1111. 1930.
[31] Taylor, H. S., & Jungers, J. C. Trans. Faraday Soc. 33: 1353. 1937.
[32] Pease, R. N., & Burnham, H. D. Jour. Am. Chem. Soc. 62: 453. 1940.
[33] Rice, F. O., & Rice, K. K. "The Aliphatic Free Radicals." The Johns Hopkins Press. Baltimore. 1935.
[34] Melville, H. W. Trans. Faraday Soc. 32: 258. 1936.

molecular split into hydrogen and acetylene in the case of mercury photosensitization, but that the 87 Kcal. excitation energy of a $^3P_1$ cadmium atom is not sufficient to produce a similar result.

To obtain further information regarding the effect of the size of the quantum on the photosensitized polymerization of ethylene, experiments were made on the reaction photosensitized by λ 2288 (*i. e.* by $^1P_1$ cadmium atoms). The $^1P_1$ state has an excitation energy of 124.4 Kcal. Up to the moment only a preliminary investigation has been made. In this a quartz lamp was used, giving λ 2288 and λ 3261. The relative intensities of the two lines were roughly 1 : 4.5 for λ 2288 as compared with λ 3261. In some experiments a Corex D filter was used to remove λ 2288. The experiments were made in a static system with metallic cadmium in the reaction vessel at 285° C. A summary of the results is given in TABLE 2.

TABLE 2

ETHYLENE POLYMERIZATION BY $^3P_1$ AND $^1P_1$

CADMIUM ATOMS

Temperature = 285° C.

Initial C_2H_4 pressure = 20 cm.

$$\frac{\text{Intensity of } \lambda\ 3261}{\text{Intensity of } \lambda\ 2288} = 4.5$$

Line	Reactant	Relative Rate
3261	C_2H_4	1
3261	$1\ C_2H_4 + 1\ H_2$	25
3261 + 2288	C_2H_4	10 5
3261 + 2288	$1\ C_2H_4 + 1\ H_2$	31 5

Whence we get the following approximate quantum yields

Line	Reactant	Approx. quantum yield
3261	C_2H_4	0 02
3261	$1\ C_2H_4 + 1\ H_2$	0 50
2288	C_2H_4	0 95
2288	$1\ C_2H_4 + 1\ H_2$	(0 59)

It is apparent that while λ 3261 is virtually ineffective, λ 2288 causes an efficient polymerization of ethylene, the quantum yield being of the same order of magnitude as that brought about by mercury photo-

sensitization with λ 2537.   It may also be noted that, as with mercury, cuprene, or some similar substance is formed in considerable quantity. This lends support to the idea that the primary step is a split into acteylene and hydrogen.   However, with cadmium there is no sign of an initial pressure increase, or of any induction period.   Also the products are mainly low boiling unsaturated hydrocarbons, especially butene.   This suggests a much closer analogy to the mechanism of the thermal polymerization than is the case with mercury photosensitization.

Some experiments have also been made with $^3P_1$ and $^1P_1$ zinc atoms (λ 3076, and λ 2139).[35]   These have excitation energies of 92.5 and 133.4 Kcal. respectively.   The experiments were made under conditions similar to those with cadmium, but at a somewhat higher temperature (355° C.) to offset the lower vapor pressure of zinc.   As yet sufficient energy measurements have not been made to enable the estimation of quantum yields, so that only relative rates are given in TABLE 3.

As yet it is not possible to estimate the relative intensities of λ 3076 and λ 2139. It appears probable, however, that the lower resonance line is relatively weaker than is the case with cadmium.   In any case, however, the similarity to the results with cadmium is striking, and it is apparent that $Zn(^3P_1)$ atoms are relatively ineffective in the polymerization of ethylene.

TABLE 3

ETHYLENE POLYMERIZATION BY $^3P_1$ AND $^1P_1$
ZINC ATOMS

			Corresponding
		Relative	Relative rate
Line	Reactant	Rate	with Cd.
	Temperature = 355° C.		
	Initial C_2H_4 pressure = 20 cm		
3076	C_2H_4	0 03	0 04
3076	$1\ C_2H_4 + 1\ H_2$	1	1
3076 + 2139	C_2H_4	0 20	0 42
3076 + 2139	$1\ C_2H_4 + 1\ H_2$	1.0	1 26

From the foregoing it follows that in the sensitized polymerization of ethylene $Cd(^3P_1)_3$ and $Zn(^3P_1)$ atoms are relatively ineffective, while $Hg(^3P_1)$, $Cd(^1P_1)$ and $Zn(^1P_1)$ atoms are effective.   Indications

[35] Steacie, E. W. R., & Habeeb, H. unpublished.

are that energy transfer in such reactions may be highly efficient, since for example, $Hg(^3P_1)$ atoms can cause dissociation of hydrogen, and Cd $(^3P_1)$ atoms can cause a C-H bond split.   As a rough but probably fairly sound approximation, therefore, we may consider that reaction will occur whenever the energy of the photoexcited atom is higher than the activation energy of the primary process.   We may therefore fix the activation energy of the reaction

$$C_2H_4 \rightarrow C_2H_2 + H_2$$

between 92.5 and 112.2 Kcal.

This is somewhat higher than the value estimated from the heat of the reaction and the activation energy of the reverse reaction (85 Kcal.).   However, both the thermochemical and kinetic data upon which this estimate was based are somewhat uncertain, and it is quite possible that the activation energy is as high as 92.5 Kcal.

There is, however, the possibility that the reaction might occur by an initial C-H bond split, as discussed previously.   If this is the case, the strength of the C-H bond in ethylene would appear to lie between 115.6 Kcal. $(Zn {}^3P_1)$ and 120.7 Kcal. $(Hg {}^3P_1)$.

Discussion

It will be apparent that the method used here to evaluate bond strengths or activation energies of primary processes is really an analogue of that suggested by Burton.[36]   Burton calculated bond strengths on the assumption that the onset of predissociation occurred when the energy of the dissociating molecule was just sufficient to rupture the pertinent bond.   The difficulty in applying this concept lies in the uncertainty in the determination of the incidence of predissociation from spectroscopic observations.[37]

The process of photosensitized decomposition may obviously be regarded as a kind of predissociation of the complex formed when the reactant molecule and the excited atom collide.

In spite of the very rough picture presented here, so many uncertainties surround existing methods of calculating bond strengths, that the method of varying the energy input by photosensitization with a variety of excited atoms seems to offer considerable promise.   In any case much interestnig information is to be anticipated from work with a variety of sensitizers.

[36] Burton, M.   Jour. Chem. Phys. **6**: 818   1938.
[37] Burton, M.   Jour. Chem. Phys **7**: 1072.   1939.

DISCUSSION

The following discussion pertaining to material in this paper was contributed by members of the conference whose names and remarks appear below.
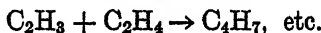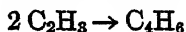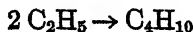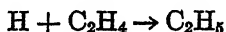
E. W. R. STEACIE:—Professor H. S. Taylor in a private discussion tells me that he has never felt satisfied with the reaction

$$C_2H_4 + Hg(^3P_1) \rightarrow C_2H_2 + H_2 + Hg(^1S_0)$$
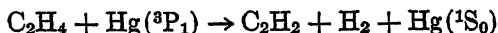
as the primary process in the mercury photosensitized polymerization of ethylene. Certainly by analogy with other reactions the step

$$C_2H_4 + Hg(^3P_1) \rightarrow C_2H_3 + H + Hg(^1S_0)$$
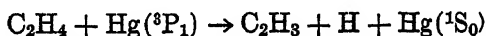
seems much more likely. The difficulty, of course, is to explain the initial pressure rise on this basis. Any set of secondary reactions such as

$$H + C_2H_4 \rightarrow C_2H_5$$

$$2\,C_2H_5 \rightarrow C_4H_{10}$$

$$2\,C_2H_3 \rightarrow C_4H_6$$

$$C_2H_3 + C_2H_4 \rightarrow C_4H_7, \text{ etc.}$$

cannot possibly lead to an initial pressure rise which persists for 20 or 30 minutes. It is perhaps possible, however, that ethylene polymerizes by some combination of both mechanisms.

H. AUSTIN TAYLOR*:—The calculation of the energy of activation of decomposition of ethylene into acetylene and hydrogen would not appear to be very well founded. The value of 42 Kcal. which van Hook and I found for acetylene hydrogenation is probably a low value. It is probable that the reaction is a chain reaction and the observed energy of activation will be complex. In any case, however, the corresponding hydrogenation in Steacie's reaction is a three body one, involving a metal atom.

Since the dissociation of Cd H is 15.5 Kcal. it would seem that the equilibrium in a system containing cadmium vapor and hydrogen atoms should be well over on the Cd H side. Hence under conditions producing low concentrations of H, those reactions which in mercury photosensitization involve H, would in the cadmium have to be attributed to Cd H. Alternatively at high H concentrations practically all the cadmium will be present as Cd H. This agrees with the observed dehydrogenation properties of cadmium as a catalyst.

---

* New York University, New York City, New York.

M. Burton*:—Dr. Steacie suggested a *maximum* value of ≈ 121 Kcal. for the strength of the C-H bond in ethylene. This value is in excess of any of the values for the C-H bond in CH_4 in current use. It seems to me that there is another value for the C-H bond in C_2H_4 which merits consideration. It has been shown recently that the C-C bond strength in acetaldehyde[1] is ≈ 75 Kcal. If we assume the same value for the single C-C bond in propylene (a not unreasonable assumption in view of the similarity of electronic configurations in the two cases) and use the generally accepted values for heats of formation, we arrive directly at a value of ∼ 93 Kcal. for the C-H bond strength in ethylene. On the other hand, if we assume that the C-C value in ethane is ∼ 3 Kcal. *less* than that in acetaldehyde, it follows that the C-H bond strength in methane is ∼ 95 Kcal., which is perhaps a minimum value.[2]

The C-H distance in CH_4 is 1.093 Å whereas it is only 1.087 Å in C_2H_4.[3] The fundamental C-H frequency is somewhat less in CH_4 than it is in C_2H_4.[4]

Consequently, it might be thought off-hand that the C-H bond strength should be higher in C_2H_4 than it is in C_2H_6, *i. e.* at least 95 Kcal. and perhaps 103 Kcal. Unfortunately, such a conclusion assumes the very point that we wish to prove by considerations such as these, *i. e.* whether we can assume similarity of anharmonicity factors in the two cases.[5] It is possible that in such polyatomic molecules the anharmonicity is greatly influenced by the electronic arrangement in the molecule, so that the C_2H_3-H link would display a greater anharmonicity than the CH_3-H link, something we might not expect by analogy to diatomic molecules.

In the discussion of Dr. Steacie's paper activation energy considerations were repeatedly employed to suggest bond strengths. It is perhaps well to emphasize that activation energy differences of forward and reverse reactions, when known from actual, unambiguous measure-

* New York University, New York City, New York.
[1] Cf. **Burton, M.** Jour. Chem. Phys. **7**: 1072. 1939.
[2] A maximum value of ∼ 103 Kcal. (equal to that for the H-H bond) is set by recent experiments on the reaction of Br and OH_4 by Kistiakowsky, who determined the activation energy of the forward process. Unfortunately, the activation energy of the reverse process could not be determined so that it was not possible to set an actual value for the C-H bond (private communication, May 7, 1940). An activation energy for the reverse process would reduce the value for the C-H bond strength by just that amount.
[3] **Pauling, L.** "The Nature of the Chemical Bond" Cornell University Press. 1939, p. 158.
[4] **Kronig, R.** "Optical Basis of the Theory of Valency" Cambridge University Press. 1935, p. 186.
[5] **Ellis, J. W.** Phys. Rev. **33**: 27. 1929 has assumed such constancy in theoretical calculations of bond strength.

ment of the reactions concerned, give true bond strengths. When such data are known there is no room for argument. Furthermore, the order of bond strengths may be suggested by the order of activation energies of homologous compounds undergoing similar reaction. In my opinion, however, it is a mistake to attempt such inference for apparently similar reactions of non-homologous compounds since in such cases the potential energy functions for the reactions concerned may be quite dissimilar.

HUGH S. TAYLOR:*—The reaction between hydrogen atoms and ethylene. The nature of the process occurring depends to some extent on the hydrogen atom concentration. With high atom concentration such as is obtained in a hydrogen discharge tube, and at low pressures of hydrocarbon, reactions of dehydrogenation predominate. The hydrocarbon molecule is both stripped of hydrogens and is ruptured to give -CH fragments. With low hydrogen atom concentrations, such as are obtained in the mercury photosensitized reaction, the system behaves as though the atomic hydrogen concentration was negligible due to rapid interaction of atoms and ethylene to form ethyl radicals. Experiment shows that more than 80% of butane is formed in such reaction conditions.

The primary process of excited mercury and ethylene. If the C-H bond strength in ethylene is greater than in ethane it may well be all the energy of an excited mercury atom would be required to produce $C_2H_3 + H$. In such case the energy of the excited cadmium atom would be insufficient. The initial pressure increase in ethylene polymerization and the known production of acetylene constitute the principal objections, as Dr. Steacie points out, to the free radical mechanism of polymerization of ethylene under the influence of excited mercury.

---

* Princeton University, Princeton, New Jersey.

PRIMARY PROCESSES IN FLUORESCENCE AND PHOTOSENSITIZATION, WITH PARTICULAR REFERENCE TO SIMPLE AROMATIC COMPOUNDS

By W. West

From New York University, New York City, New York

Chemical photosensitization, the rendering chemically active of an otherwise inert substance by light absorbed by another material in the system, seems to have essentially two types of mechanism. One may be represented by the scheme $A + S^* = A^* + S$, in which, subject to energetic and quantum-mechanical restrictions, a transfer of energy occurs from the excited sensitizer S^* to the reactant A. This is activated to the condition A^*, here corresponding to a chemically altered state, while the sensitizer reverts to its unexcited state. The transfer may occur at collision, as in the best understood cases, such as sensitized fluorescence in gaseous atoms and dissociation of simple molecules by activated atoms; but also, by the formation of coordination or adsorption compounds between A and S, the sensitizing and accepting parts may be permanently associated, and then the process verges towards that type of photochemical reaction in which energy is acquired by one part of the molecule and transferred within the molecule to another bond whose disruption initiates chemical change. Photosensitization by the uranyl ion in the decomposition of oxalic acid, and the sensitization of the photographic plate by dyes are probably examples of the latter process.

The other type of mechanism that has been recognized in photosensitization involves the production by the action of light on the sensitizer of entities which may participate in a chain reaction, as in the production of water from hydrogen and oxygen by visible light in the presence of chlorine.

$$Cl_2 + h\nu = Cl + Cl$$

$$Cl + H_2 = HCl + H$$

$$H + O_2 = HO_2, \text{ etc.},$$
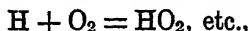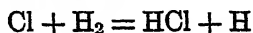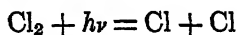
the eventual products being HCl and H_2O. We shall not have to discuss this type of sensitization.

The interest of the sensitized reactions to be discussed in this paper is that they seem to represent, in solution, and with fairly complex molecules, the simplest type of collisional activation.

In the course of an investigation of the photodecomposition of ethyl iodide in solution,[1] the observation was made that the rate of decomposition in benzene, in which the bulk of the light was absorbed by the solvent, was as great as in hexane, which is almost completely transparent to the effective radiation.   Not all absorbing solvents behave in this way, and the decomposition of ethyl iodide in benzene must therefore be regarded as a photosensitization, $B^* + EtI = B + De$-composition Products, even though, owing to the approximate identity in position of the absorption spectra of the iodide and of benzene, the sensitizer will not bring about the reaction at wavelengths other than those effective in direct absorption by the iodide.   Naphthalene, however, has an absorption spectrum somewhat similar to that of benzene, but at longer wave lengths, in regions of feeble absorption by ethyl iodide, and it turns out that in these regions the presence of naphthalene markedly accelerates the photodecomposition of ethyl iodide. Many of the experiments have been carried out in collaboration with my colleague, Mr. Walter Miller, of whose assistance I should like to make due acknowledgement.   Some of the quantitative results have appeared in reference 2 which also contains details on experimental methods.

It may be said at once that the fact that the absorption spectra of naphthalene and of ethyl iodide are mutually unaffected in mixtures of the two dissolved in hexane is very strong evidence that no long-lived complex between these molecules exists.   The sensitization must therefore occur at collision.   Naturally, the conclusion is not drawn that at these collisions the energy levels of the partners are not mutually influenced; the small population of molecules in the state of collision would make such effects on the absorption inappreciable by the method of measurement that was used.

FLUORESCENCE AND SENSITIZATION

It is well known that in the simplest cases of gaseous photosensitization, the sensitizer, in the absence of the acceptor, emits its excitation energy as fluorescence, and that this emission is quenched when the acceptor is added.   In this case, sensitization and fluorescence are

[1] West, W., & Paul, B.   Trans. Faraday Soc. 28: 688.   1932.
[2] West, W., & Miller, W.   Jour. Chem. Phys. 8: 849.   1940.

simply alternative processes by which the excited molecule reverts to its lowest state. In the liquid state, matters are not so simple. Neither are all sensitizers fluorescent in the absence of the acceptor (*e. g.* in Eder's reaction, the photo-reduction of mercuric to mercurous chloride by oxalic acid sensitized to visible light by ferric ion) nor is intensity of fluorescence within a group of sensitizers a measure of their sensitizing efficiency (for instance, Eder's reaction is sensitized more effectively by the feebly fluorescing erythrosine than by fluorescein). On the other hand, it appears obvious that not fluorescence, but the quenching of fluorescence, might be expected to be a measure of sensitization efficiency towards the quencher. A fluorescence substance is always a potential sensitizer, and fluorescence and photosensitization must always have something of the nature of alternative processes, though in the case of complicated molecules, not necessarily of simple alternatives. These almost self-evident relations seem to be rather curiously obscured in much of the older literature on the subject, perhaps because the reactions studied, such as Eder's reaction, a heterogeneous chain reaction, have important features besides the transfer of energy. The recent work of Bowen on the photo-oxidation and fluorescence of rubrene[3] points to a simple reciprocal relation between fluorescence and chemical activation of the fluorescent molecule in the proper kind of environment, and a similar parallelism is found in our experiments between the capacity of a fluorescent substance to be quenched and to sensitize the decomposition of the quencher. It seems therefore appropriate to devote some attention to the salient points on the fluorescence of naphthalene and its derivatives, and its quenching, which turn out to be quite typical of the whole of these phenomena in solution.

FLUORESCENT PHENOMENA IN NAPHTHALENE VAPOR AND SOLUTION

The absorption spectra of naphthalene and its derivatives have been carefully measured by Victor Henri and H. de Laszlo.[4] Between about 3200 and 2900 Å a series of well defined sharp-headed bands is observed in naphthalene, with maximum extinction coefficients of about 100, and with the appearance of bands associated wtih a transition between two stable electronic states. This is followed between

[3] Bowen, E. J. Trans. Faraday Soc. 35: 15. 1939.
[4] Henri, V., & de Laszlo, H. Proc. Roy. Soc. A105: 662. 1924. De Laszlo, H. Proc. Roy. Soc. A111: 355. 1926.

2800 and 2500 Å, (FIGURE 1) by a region of much more intense absorption ($\epsilon = 2000$ to 5000) in which the maxima have the diffuse appearance of predissociation spectra. This type of absorption is observed both in the vapor and in solution, the main difference being a relatively small shift of the bands to longer wave length in solution. The derivatives also show the same two regions, though the intensity of absorption is usually greatly increased over that of naphthalene in the long wave region, and sometimes both regions are diffuse.
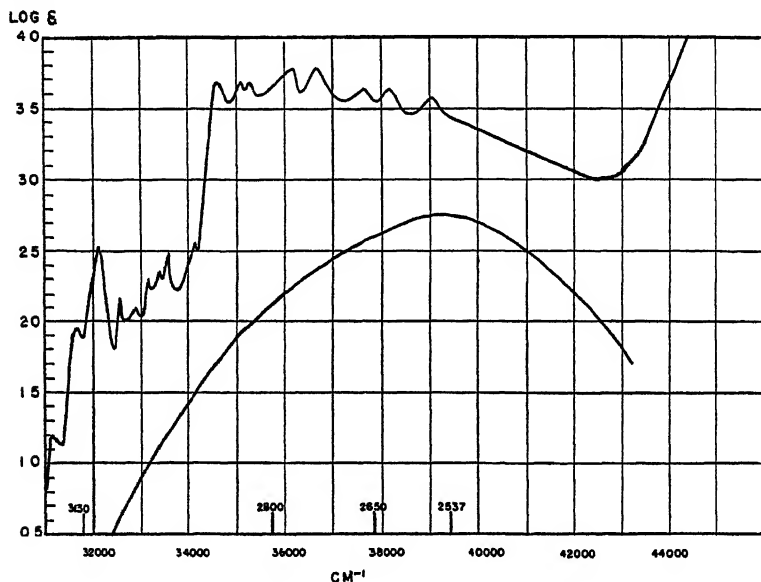
FIGURE 1   Absorption of naphthalene and of ethyl iodide.

Fluorescence in solution in naphthalene occurs as a series of bands on the long wave side of the first absorption region, and has the same appearance, and the same quantum efficiency, for all exciting wave lengths, whether in the region of well defined or diffuse bands. In the absorption of light by naphthalene in the near ultraviolet, three electronic levels seem to be concerned (FIGURE 2): A, the lowest; B, the upper level for the sharp vibrational bands; and C, the upper level for the diffuse bands. The independence of the fluorescence spectrum in solution on the absorption band used for excitation points to there being a common initial state for the fluorescent emission, irrespective of the vibrational level attained as a result of absorption. A molecule excited by absorption to a high vibrational state of

electronic level B is deactivated to the lowest vibrational state of that level. It is the vibrationless state of B that resists collisional deactivation, internal utilization of the energy, and in general all processes that prevent radiation. On the other hand the nature of electronic level C is such that deactivation from C to B is practically certain to occur, so that excitation to a level in C produces the same ultraviolet fluorescence as to one in B.

The appearance of fluorescence by excitation in a region of "predissociation," as represented by the A → C absorption spectrum, has important photochemical implications. The phenomenon has already
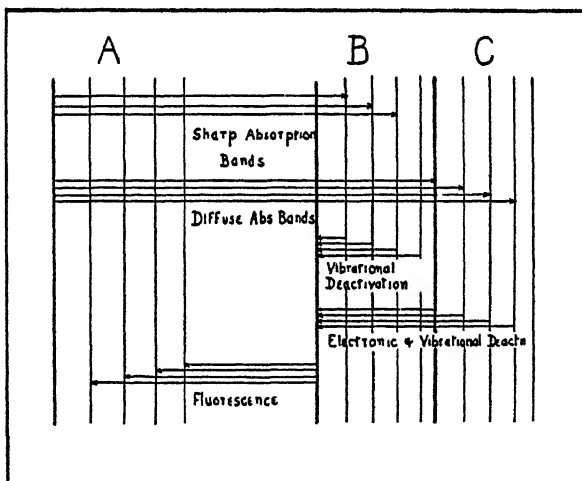
FIGURE 2   Energy levels in naphthalene (schematic).

been observed and commented on by Terenin, Vartanian, and Neposent,[5] who found that aniline fluorescence continues with undiminished yield in hexane solution well within the predissociation region of absorption, while the vapor shows a sharp drop in yield when the exciting radiation passes from the sharp to the diffuse absorption region.

In naphthalene vapor, we have found that excitation in both regions produces fluorescence on the long wave side of the region of sharp absorption, even at pressures of 0.01 millimeter. But the emission is no longer identical for the two regions of excitation,—whereas excitation in the sharp region produces sharp-headed fluorescent bands,

⁵ Terenin, A., Vartanian, A., & Neposent, B.   Trans. Faraday Soc. **35**: 39   1939

excitation in the diffuse region produces a continuous emission, in the same general region as the other, but extending to somewhat shorter wave lengths, overlaid by very diffuse maxima, and moreover, as in the aniline case, the quantum yield of fluorescence excited by wave length 2537Å, in the diffuse region, is only about one-third of that excited by absorption in the sharp region.

The occurrence of the B → A fluorescence with A → C excitation at pressures as low as 0.01 millimeter, shows that either the C → B transition occurs internally in the individual molecule, or, if at collision, with much greater than kinetic cross-section. In the former case a separate C → B fluorescence, at about 12,000 Å is to be expected. At this region both photographic plates and photoelectric cells are of low sensitivity, and the thermopile is in general not very sensitive to low intensities. Our endeavors to detect radiation at this region in the fluorescence of naphthalene have failed, but in view of the difficulties of detecting feeble radiation of this wave length, are inconclusive.

The difference in character between the fluorescence of naphthalene in solution and the vapor can be easily explained in terms of FIGURE 2. In the low pressure gas, vibrational deactivation can no longer take place by collision, so that a large number of vibrational levels of B will be occupied after the transition C → B occurs. The fluorescence spectrum originating in the electronic state B, now, in contrast to the situation in solution, distributed over many vibrational levels, will be correspondingly rich in bands, which by overlapping could give the appearance of a continuum at low resolution.

In concluding this discussion of the difference in the variation of fluorescence efficiency with exciting wave length shown by aniline and naphthalene in the dissolved and gaseous states we may point out its general significance. In polyatomic molecules in solution, excitation in a region of absorption which in the corresponding gas has all the characteristics of a predissociation spectrum need not lead to permanent disruption of the molecule. Here seems to be a very direct proof of the reality of the process of "primary recombination" of disrupted fragments in solution described by Franck and Rabinowitch.[6]

• Franck, J., & Rabinowitch, E.   Trans. Faraday Soc. 30: 125.   1934.

FLUORESCENCE OF DIFFERENT NAPHTHALENE DERIVATIVES IN SOLUTION

Naphthalene derivatives in solution show the same characteristics in their fluorescence as the parent substance,—a spectrum independent in appearance and in quantum efficiency of the exciting wave length. The value of their quantum efficiencies differ greatly from compound to compound, as do also the character (band or continuum) and to a smaller extent the wave length regions of the fluorescence. Approximate relative quantum efficiencies for a number of compounds are given in TABLE 1.

TABLE 1

RELATIVE FLUORESCENCE EFFICIENCIES OF NAPHTHALENE COMPOUNDS

Derivative	α-Chlor	Naphthalene	α-COOH	β-CN	β-OH	β-NH₂	α-NH₂
Efficiency	.05	.15	.3	.5	.75	1	1
Character of First absorption region	sharp bands	sharp bands	fairly sharp bands	fairly sharp bands	very diffuse bands	fairly sharp bands	practically continuous

These numbers are also probably not very far from the absolute efficiencies of fluorescence; Bowen[3] gives .2 as the absolute value for naphthalene. Not all naphthalene derivatives fluoresce; among others, the nitro-compounds do not, and one would expect, on analogy with the corresponding benzene compounds, that bromnaphthalene would fluoresce only slightly, and the iodo-compounds not at all.

QUENCHING OF FLUORESCENCE

The complete or partial extinction of fluorescence in solution by the addition of certain substances seems to have been discovered by Stokes[7] in his pioneering studies on fluorescence. On analogy with the quenching of gaseous fluorescence, so thoroughly studied by Franck and his colleagues, it has seemed very probable that quenching in solution was likewise in essence due to resonative transfers of electronic energy between the fluorescing molecule and the quencher. Franck and Levi[8] have especially focussed attention on the conditions that must be fulfilled to ensure a high probability of such a transfer: as complete as possible conversion of all the electronic energy of the fluorescent molecule into electronic energy of the acceptor, as distinct

[7] Stokes, G. G. Phil. Trans. Roy. Soc. **142**: 541. 1852.
[8] Franck, J., & Levi, H. Zeit. Phys. Chem. **B27**: 409. 1934.

from translational or other thermal energy. (The kinetic energy of separation of fragments of the acceptor, if it is disrupted after the transfer, is irrelevant.) For example, the fluorescence of naphthalene is very effectively extinguished by ethyl iodide, and the mechanism seems to be essentially a resonative transfer at collision, subject to the rules of collisions of the second kind, in which the electronic energy of the naphthalene raises the ethyl iodide to a repulsive electronic state, at about the same energy level as the excited naphthalene, leading to a dissociation of the iodide.

In addition to transfer of energy at collision to an external body, a kind of "internal quenching" appears to occur, and indeed to be the main reason for the absence of fluorescence in substituted fluorophores like the halogenated or nitrated aromatic compounds. At least two cases might arise. A repulsive energy surface involving, say, the carbon-halogen distance might come close to the surface representing the upper state concerned in fluorescence, in which case quenching by predissociation could occur, with the disruption of the carbon halogen bond. Another consideration is suggested by the work of Sklar and his co-workers[9] on the spectrum of benzene. In this compound, the pure electronic transition A → B (FIGURE 1) is forbidden for reasons of symmetry, the prohibition being somewhat weakly evaded in the hydrocarbon by a simultaneous excitation of certain types of vibration of the appropriate symmetry, and also, in substituted molecules, by asymmetries introduced by the substituent. Any flux of energy within the excited molecule to vibrational modes whose symmetry forbade a radiational transition to the normal state would cause quenching. On the other hand it does not follow that such an internal energy flux necessarily deactivates the excited molecule as a photosensitizer in a collision process, and here we may perceive one reason for the association in certain substances of high sensitizing ability with feeble fluorescence.

There has been some discussion, most recently by G. N. Lewis and M. Calvin,[10] of a possible relation between high fluorescence efficiency and sharp absorption bands, and conversely between low fluorescence and diffuse absorption. Unfortunately for the application of this as a simple criterion for fluorescence, no such general relation exists; TABLE 1 for instance, shows cases of practically continuous absorption associated with high fluorescence, and of relatively

[9] Sklar, A. L. Jour. Chem. Phys. 5: 669. 1937. 7: 984. 1939. Sponer, H., Nordheim, G., Sklar, A. L., & Teller, E. Jour. Chem. Phys. 7: 207. 1939.
[10] Lewis, G. N., & Calvin, M. Chem. Rev. 25: 273. 1939.

sharp absorption with low fluorescence. It is quite true that diffuseness in a vibrational band indicates a low duration of the excited state, which, if not due to dissociation, might be due to energy flux within the molecule to other stable vibrational modes, but the empirical occurrence of intense fluorescence with diffuse absorption shows that the mere occurrence of the flux does not prevent fluorescence: apart from dissociative processes, flux to modes incapable of combining with the normal state would be required to suppress fluorescence. But if the degree of sharpness of absorption bands cannot be accepted indiscriminately as a criterion of fluorescence power, evidence appears to be accumulating that there is indeed some intimate relation between these qualities, subject to strong influence by other factors, as perhaps that suggested above. Lewis and Calvin[10] have suggested a relation between structural rigidity in the molecule and the appearance of sharp absorption bands and strong fluorescence. They give several examples of these correlations, and, under the influence of their suggestions, I have found sharpness of absorption to go parallel with intensity of fluorescence in some simple benzene derivatives, *e. g.* in the three xylenes.

EXCITATION TIMES IN FLUORESCENT MOLECULES

An important property of an excited molecule is the time τ it remains in the excited condition before it radiates. Other factors being equal, τ will determine the efficiency of a molecule as a fluorescer or as a sensitizer. For the greater the average period of existence of the excited molecule before radiation, the greater the probability of energy transfer by internal or external quenching and sensitization, and the feebler the fluorescence. In a group of related molecules with absorption in the same general spectral region, the best fluorescers will have the shortest τ, and energetic and quantum conditions for transfer being equally well fulfilled, the best sensitizers will have the longest τ.

The methods available for the determination of average lives have been recently reviewed by Hirschlaff,[11] and I shall here merely recount two methods relevant to the subsequent discussion.

From Intensity of Absorption

Defining the absorption coefficient k by the relation $I_\nu = I_{0\nu}\, e^{-k_\nu d}$ where $I_{0\nu}$ and I_ν are respectively the incident intensity of radiation

[11] Hirschlaff, E. "Fluorescence and Phosphorescence." Methuen & Co. London. 1938.

of frequency ν sec^{-1} and that transmitted by a layer of absorbing medium d cm. thick, one can obtain the relation

$$\int k_\nu d\nu = \frac{\lambda_0}{8\pi} \cdot \frac{g_2}{g_1} \cdot \frac{N}{\tau}.$$

The integration is to be carried over the frequency range covered by the electronic transition in question, λ_0 is the wave length of the centre of the band, N the number of unexcited atoms per cubic centimeter, and g_2 and g_1 the statistical weights of the excited and unexcited electronic states. The equation is valid for the electronic transitions of atoms[12] and neglecting coupling between the rotational and vibrational motions and the electronic, should be applicable to the electronic band systems of polyatomic molecules, though in practice the evaluation of g_2 will demand a greater knowledge of the excited state than is often available at present.

Methods Depending on the Quenching of Fluorescence

Theories of quenching are nearly all based on the idealization

$$F + h\nu \rightarrow F^* \qquad\qquad (1)\ \text{absorption}$$

$$F^* \rightarrow F + h\nu \qquad\qquad (2)\ \text{fluorescence}$$

$$F^* + X \rightarrow F + X^* \qquad\qquad (3)\ \text{quenching}$$

The simplest theory, first given by Stern and Volmer,[13] reduces to $I/I_0 = 1(1 + kc)$ where I and I_0 are the quenched and unquenched intensities of fluorescence respectively, c the concentration of the quencher, and k a constant involving τ and the interval between collisions. A corollary of this theory, whose applicability is carefully discussed by Mitchell and Zemansky,[12] is that τ equals the average interval between collisions of the excited molecule and quenching molecules when the latter are present at such a concentration as to diminish the fluorescence by half, provided every collision is a quenching one. It will be seen that the quenching formula contains two unknowns, the average life of the excited fluorescent molecule, and the effective diameter for the quenching collision, the latter of which cannot usually be independently evaluated.

[12] See, for instance, Mitchell, A. C. G., & Zemansky, M. W. "Resonance Radiation and Excited Atoms." Cambridge University Press. 1934.
[13] Stern, O., & Volmer, M. Phys. Zeit. 20: 183. 1919.

EXPERIMENTAL OBSERVATIONS OF QUENCHING IN SOME NAPHTHALENE COMPOUNDS

TABLE 2, and FIGURES 3 and 4 contain data on the quenching of the fluorescence of naphthalene and some of its derivatives by ethyl iodide,—the exciting radiation was of wave length 3130 Å. The

TABLE 2

QUENCHING OF NAPHTHALENE COMPOUNDS BY ETHYL IODIDE: VALUES OF $\dfrac{I}{I_0} \times 100$

Fluorescing substance	conc. of EtI in millemoles/l							Quenching constant (millemole^{-1})
	1.24	3.72	12.4	19	24.8	49.6	95.5	
Naphthalene 1.9×10^{-3}m in EtOH ($\eta = .01$)	83.6	61.0	20					.176
Naphthalene 1.9×10^{-3}m in EtOH- glycerine 1/1	95		70	54				.040
β Naphthol 1.9×10^{-4} in EtOH	95.8	91.1	86.9	78.2	74 8	54	33.9	.014
α Naphothoic acid, 2.67×10^{-5} in EtOH				98.8		95		.00
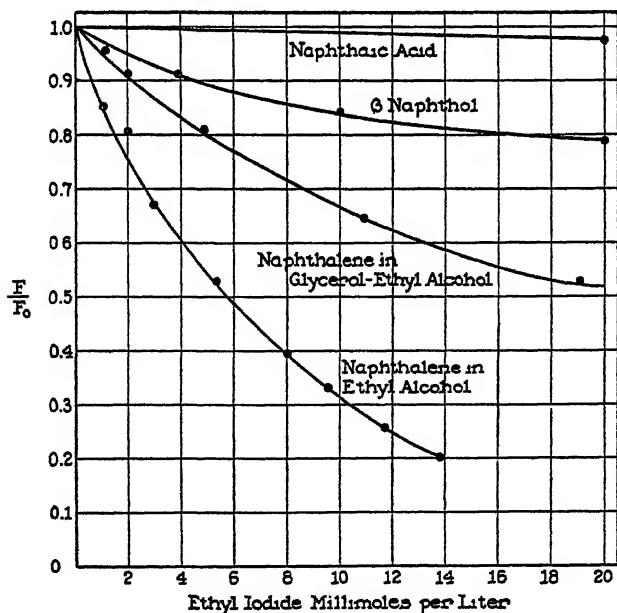
FIGURE 3. Quenching of fluorescence by ethyl iodide.

quenching constants in this table are the values of k in the Stern-Volmer expression, $i.$ $e.$ the slope of the graph of I_0/I as a function of the concentration of quencher.

The following characteristics appear: a linear relation between I_0/I and concentration of quencher at low values of these concentrations; inhibition of quenching by an increase in viscosity of the medium; unequal efficiencies of a given quencher with respect to different fluorescent molecules, even although chemically related and showing
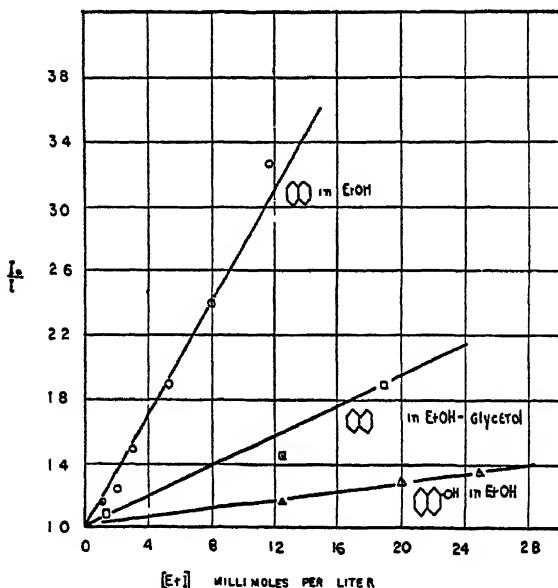

FIGURE 4. Quenching of fluorescence of naphthalene compounds by ethyl iodide

absorption and fluorescence in the same spectral regions. Whereas an addition of about .005 moles per litre of ethyl iodides reduces the fluorescence of naphthalene to half its value, ten times as much is required to do the same in β naphthol, and α-naphthoic acid is affected only at very high concentration of quencher. α chlornaphthalene is likewise little quenched by ethyl iodide, and incidentally, in view of the great efficiency of iodide ion in suppressing the fluorescence of certain fluorescent ions, as quinine bisulphate, uranyl, fluorescein, and other dyes, it may be mentioned that it has little effect on the fluorescence of naphthalene or its derivatives.

MECHANISM OF QUENCHING AND PHOTOSENSITIZATION

The effect of an increase in viscosity in suppressing quenching, and, as we have found experimentally, sensitization, shows definitely that the processes take place at collision involving the diffusion of the quencher to the excited molecule. The precise way in which the viscosity of the medium influences collision processes in solution has been very clearly demonstrated by E. Rabinowitch.[14] The viscosity of the medium does not affect the total number of collisions between solute molecules in a solution, averaged over a long period. But whereas, within obvious limits, the short time average of collision frequency in gases is the same as the long time average, this is not true for solute molecules in liquid solution. Here the collisions appear to take the form of sequences of rapid impacts, having something of the nature of vibrations, between the colliding molecules enclosed in a cage of solvent molecules, separated by relatively long intervals during which the solute molecules diffuse through the medium until a new encounter takes place which can allow a new sequence of collisional vibrations to occur. The number of impacts in a group is increased by increase of viscosity, but the time between groups correspondingly decreased, with the result that the long time average of encounters is independent of viscosity. The frequency of events, such as chemical reactions requiring energy of activation, which occur only rarely at collision, will not be appreciably influenced by viscosity. But processes that occur at the *first* impact between the reacting molecules will be slowed in viscous media, since the frequency of these events is determined by the interval between collision groups. Since quenching and photosensitization in solution are inhibited in viscous media, the conclusion follows that these processes are very efficient and take place at the first collision between the reactants.

Some remarks may be made on the electronic process occurring at the collision. Studies of the absorption spectrum of ethyl iodide have shown that in the spectral region in question the electron primarily excited is a non-bonding one attached to the iodine atom,[15] but there is little doubt that excitation of this electron is followed practically immediately by disruption of the C-I bond. The process attending the direct absorption of light by ethyl iodide can then be taken as $EtI + h\nu = Et + I$. This takes place with highest probability at

[14] Rabinowitch, E., & Wood, W. C. Trans. Faraday Soc **32**: 1381. 1936.
[15] Mulliken, R. S. Phys. Rev. **47**: 413. 1935.

about 2500 Å, the maximum of the absorption spectrum, and is relatively improbable at 3130 Å, where the intensity of absorption is low. We have to explain how the energy in a photon of wave length 3130 Å, taken up from radiation with such low probability, is avidly accepted from a naphthalene molecule possessing just the same amount of energy.

The energy relations are depicted approximately in FIGURE 5. (Weiss and also Rabinowitch[16] have pointed out that the non-adiabatic nature of resonance transfer of the type under discussion

C₂H₅ . . . I DISTANCE.
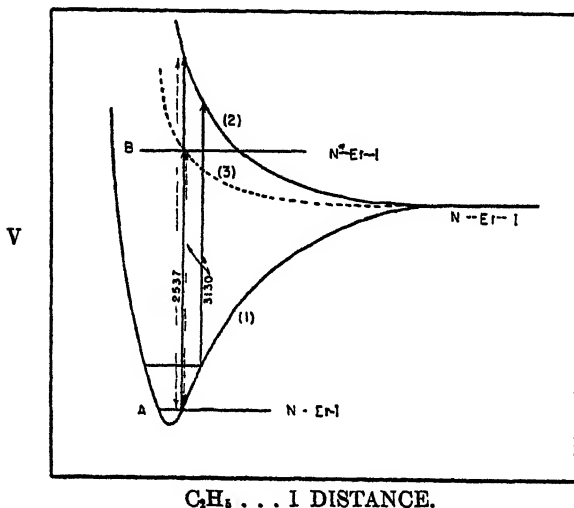
FIGURE 5. Mechanism of photosensitization.

precludes in principle the drawing of potential energy curves as a function of the distance apart of the interacting centres, since as the interaction depends on time factors, no unique energy,—distance correlation is possible, but for illustrative purposes, probably no great harm is done in using the curves.)   Curve (1) is the potential energy curve for various Et . . . I distances in the normal electronic state with a naphthalene molecule, *N*, at an infinite distance off. As the interaction between normal naphthalene and normal ethyl iodide is small, this curve will be only slightly changed when the naphthalene is brought up to near the distance corresponding to a collision *i. e.* the line *A* represents approximately the energy N . . Et . . I at collision.

16 Weiss, J. Trans. Faraday Soc. 35: 48. 1939. Rabinowitch, E. Trans. Faraday Soc. 35: 61   1939.

Curve (2) is the repulsive state of EtI, responsible for its dissociation, with the normal naphthalene at infinity. The ethyl iodide molecules in their lower vibrational states absorb about 2500 Å, and only vibrationally excited ones at 3130 Å. An electronically excited naphthalene molecule, however, exerts attractive forces on ethyl iodide molecules and at the collision distance between such a naphthalene and an ethyl iodide molecule at which resonative transfer takes place, curve (2) is lowered to (3) the line B representing the energy N^* . . Et . . I at collision. $A \to B$ represents the electronic transition already referred to as associated with the fluorescence, as modified by the colliding EtI, which the naphthalene can undergo by absorption of 3130 Å. The principle of maximum utilization of electronic energy is fulfilled if transfer occurs at this point, and we can see how vibrationally unexcited ethyl iodide in the process of collision with

TABLE 3

	Naph-thalene	α-Cl	β-ON	β-OH	α-COOH	β-NH₂	α-NH₂
$\tau = \frac{g_2}{g_1}$ (sec)	1.06 ×10⁻⁶	1.6×10⁻⁷	9.9 ×10⁻⁸	4.3 ×10⁻⁸	1.7×10⁻⁸	8.3 ×10⁻⁹	2 3 ×10⁻⁸
Quenching Constant for EtI	.176	low		.014	0 00	Chemical reaction	
Relative Fluorescence Efficiency	.15	.05	.5	.75	.3	1	1
Photo-sensn. Efficiency to EtI	high	none	<naph.	<naph.	none	—	—

the excited naphthalene can now utilize the energy in a photon of wave length 3130 Å and dissociate.

What is the origin of the different tendencies of different naphthalene compounds to be quenched in fluorescence and to sensitize a given reaction? The energy available in the different compounds studied is about the same: the energy of excitation is not the vital factor. Differences in probability of transfer could originate through different interactions between the excited molecules and ethyl iodide displacing the intersection of the dotted curve of FIGURE 6 with the line B from the favoiable position shown: or through differences in the duration of the activated state, besides through specific interactions between substituents in the excited molecule and the quencher. If different effects on the energy levels were responsible for the differences in quenching and sensitization, we should be able to find other wave lengths at which the substituents would sensitize well; there is no evidence that this is so. The information we have at present scarcely

enables us to determine which factor is predominant in determining sensitization efficiency: but there does seem a significant parallelism between the active life of the excited compound, as determined by the intensity method (no direct determinations being available) and its resistance to quenching. (TABLE 3.)

The exceptional position of chlor-naphthalene, probably due to strong internal quenching, has already been commented on.

DATA ON SENSITIZATION EXPERIMENTS

The primary process in the direct photolysis of ethyl iodide is very probably, as we have already stated, the dissociation $EtI + h\nu = Et + I$. The observed quantum yield, in terms of the number of iodine atoms produced per photon absorbed is always less than unity, about .3 at wave length 3130 Å and .4 at 2600 Å in hexane solution. These yields are independent of the concentration of iodide down to values below .001 m., and of light intensity, and their deficiency from unity seems to be due largely to recombination of the radical and the iodine atom.

When naphthalene is added to a dilute solution of ethyl iodide in hexane, there is a much increased absorption of light by the solution in the neighborhood of wave length 3100 Å. For instance the addition of 2×10^{-3} molar naphthalene to 5×10^{-3} molar iodide increases the absorption about 10 times, and the rate of decomposition of the iodide is increased in the same proportion. The light absorbed by the naphthalene is available for the decomposition of the iodide, which decomposes at the rate at which it would have if all the light had been absorbed by itself.

Effects of Concentration

The over-all quantum yield, that is the ratio of the number of iodine atoms produced to the total number of photons absorbed is found to be independent of the naphthalene concentration; at very low concentrations of naphthalene, the bulk of the light is absorbed by the iodide, which decomposes by the unsensitized reaction; at higher naphthalene concentrations the sensitizer absorbs most of the light and the iodide decomposition is sensitized at about 3100 Å. The over-all yields in the direct and sensitized decompositions are the same, showing that the fate of the ethyl iodide molecule is the same whether it has absorbed energy from a photon of wave length 3100 Å or from a naphthalene molecule excited by the absorption of such a photon,

both with respect to the primary dissociation and the secondary reactions which determine the yield.

On the other hand, the sensitized yield of ethyl iodide falls off from the normal value at ethyl iodide concentrations below about .01 m., in marked contrast to that in the direct photolysis. These effects of concentration of sensitizer and sensitized reactant are illustrated by TABLES 4 and 5, and by FIGURE 6.

TABLE 4

EFFECT OF NAPHTHALENE CONCENTRATION

[EtI] = .02 molar. λ 3130 Å; Fraction of light absorbed by EtI without naphthalene .25

[Naph.]	.6 molar	8×10^{-2}	1.8×10^{-2}	4×10^{-3}	1.8×10^{-3}	9×10^{-4}	1.6×10^{-4}
Fraction of light absorbed in sensitized reaction	1	1	1	.73	.60	.57	.35
Quantum Yield	.32	.33	.31	.28	.31	.32	.32

TABLE 5

EFFECT OF ETHYL IODIDE CONCENTRATION

[Naph.] = .0018 m. λ 3130 Å; Fraction of light absorbed by naphthalene .55

[EtI]	2.5×10^{-4}	2.1×10^{-4}	1.8×10^{-3}	5×10^{-3}	1.4×10^{-2}	2.1×10^{-2}	8×10^{-2}
Yield	.051	.075	.155	.280	.33	.31	.33
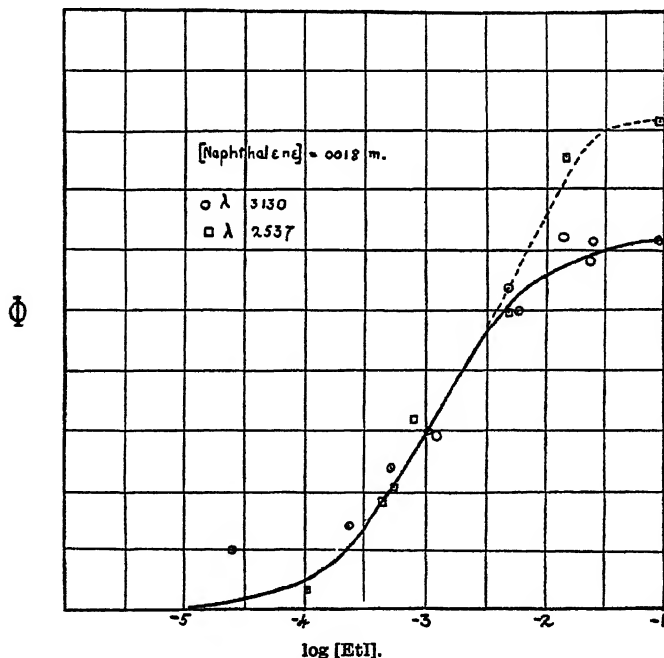
FIGURE 6. Log [ethyl iodide].

Effect of Viscosity

Like the fluorescence quenching, the sensitizing effect of naphthalene is materially diminished by an increase in the viscosity of the medium. Substitution of nujol for hexane as solvent reduces the quantum yield for decomposition of ethyl iodide to about half its value in hexane.

Effect of Wave Length

At wave length 3130 Å, the sensitized and unsensitized reactions have the same yield; .3 atoms of iodine produced per quantum absorbed. At wave length 2537 Å, direct yield is about .4, but the sensitized is practically the same as at wave length 3130 (FIGURE 6). This interesting fact can be understood in terms of our previous discussion of the details of the fluorescence of naphthalene and of the mechanism of transfer. Absorption of wave length 2537 Å raises the naphthalene molecule to a high vibrational state in the electronic level C (FIGURE 2) which was shown by the fluorescence spectrum to be very rapidly deactivated to the vibrationless state of the lower electronic level B. This deactivation occurs on the average before an effective collision can be made with an ethyl iodide molecule; only the residual electronic energy of state B is available for transfer, and the sensitized yield, like the fluorescence yield, is independent of the exciting wave length. At very high iodide concentrations, the efficiency of the sensitized reaction at wave length 2537 Å, might be expected to increase, a matter for further investigation. However, even if deactivation as described did not occur, the defect from resonance that exists in this situation would cause a diminution in the probability of transfer. The energy corresponding to wave length 2537 Å is considerably above that required to carry a normal ethyl iodide molecule to the repulsive curve by a "vertical transition" and a fraction of the excitation energy would have to be expended in kinetic energy, leading in accordance with the energetic principle of transfer, to a diminished probability of transfer.

ACTIVE LIFE OF SENSITIZER FROM SENSITIZATION MEASUREMENTS

As shown by the data of TABLE 5, there is in the sensitized reaction, in marked contrast to the direct photolysis, a notable fall in the quantum yield at low ethyl iodide concentrations. According to the collision mechanism which has been adduced for the sensitized reaction, such a fall must occur at concentrations at which the interval

between collisions of an excited naphthalene molecule becomes comparable to the active life of the former; above these concentrations, most of the naphthalene molecules are deactivated by energy transfer, below, to an increasingly greater extent by radiation. The value of the ethyl iodide concentration at which the sensitized yield begins to fall off from the high concentration value is of the same order as that required to reduce the fluorescence of naphthalene by half, but is numerically somewhat greater, a relationship which analysis has shown indicates that the sensitizing reaction can draw on excited naphthalene molecules other than those which would fluoresce in the absence of the ethyl iodide.[2]

An estimate of the active life of the excited naphthalene molecule should then be obtainable from a knowledge of the concentration at which the sensitized yield deviates from the limiting value, if we know the appropriate collision diameter and the law for the frequency of collisions. If we use for the diameters the approximate kinetic values of 5×10^{-8} cm. and the simple gas-kinetic expression for collision frequency, the period between collisions of ethyl iodide and naphthalene at the critical concentration turns out to be about 10^{-9} sec. This is in serious disagreement with the value of $\tau \times g_2/g_1 \doteq 10^{-6}$ sec., for naphthalene, calculated from intensity measurements, a disagreement that cannot be removed by a consideration of the statistical weights, whose ratio g_2/g_1 is probably greater than unity. If the long value of 10^{-6} sec. for the active life of the excited naphthalene is even approximately correct, as seems reasonable in view of the forbidden nature of the transition, we might conclude from the two discordant values that the effective collision diameter for sensitization is much smaller than the kinetic (some .03 times the kinetic) and proceed to seek in steric factors a reason for the apparent low collision efficiency. While it seems plausible that steric factors must be of greater importance in transfer of energy between polyatomic molecules than in the atomic case, it was pointed out by Dr. E. Rabinowitch in his discussion of this paper at the symposium that the essential reason for the discrepancy lies in the use of the gas collision formula for a collision process in solution which occurs in the manner already indicated in the discussion of the effect of viscosity on quenching and sensitization. Very efficient processes which occur at the first encounter in a collision group will appear inefficient in terms of gas kinetic theory, since all the other collisions of the group will be useless for the process. The average time between collision groups in the particular case under discussion has to be compared with the life-

time of the excited sensitizer, and the ratio of the true life-period to that calculated from gas kinetic theory will be some measure of the number of collisions per collision group.

SENSITIZATION BY OTHER COMPOUNDS

TABLE 6 contains a list of the substances we have examined, arranged roughly in the order of their sensitizing action towards ethyl iodide.

TABLE 6

Best sensitizers	Less efficient	Non-sensitizers
naphthalene	α-naphthol	α-chlornaphthalene
acenaphthene	β-naphthol	α-nitronaphthalene
fluorene	α-naphthyl ethyl ether	α-naphthoic acid
	β-naphthyl ethyl ether	anthracene
	α-naphthylamine 4-sulphonic acid	
	1-nitro 2-naphthylamine	
	α-naphthonitrile	
	β-naphthonitrile	

The hydrocarbons appear to be the best sensitizers. We have already briefly discussed one general effect of substitution which might tend to lower sensitization and increase fluorescence, namely the effect of a substituent in shortening the life of the excited state by removing the restrictions on a radiational transition from the excited to the normal electronic level. But obviously specific properties of the substituent are of importance, and interactions between different substituents in the same molecule. (For example, the inclusion of 1-nitro 2-naphthylamine among the sensitizers in TABLE 6 shows that the introduction of the amino group has counteracted the inhibition of sensitization and fluorescence caused by the nitro group.) Detailed understanding of the effect of substituents must await further experimental work.

SUMMARY AND GENERAL CONCLUSION

The essential conclusion of the experiments reported here is that they demonstrate the existence, among polyatomic molecules in solution, of dissociative chemical reactions induced by the transfer of electronic energy from one constituent of the reaction system to another at collision. The transfer may be very efficient, and its characteristics seem to differ from those of the corresponding resonance

transfers hitherto studied in gases of simple molecular constitution only in modifications which can be rationally associated with the polyatomic nature of the reactants, and with the nature of collisions in liquids.

The following points may be emphasized in connection with resonance transfers between molecules of some complexity in liquid solution.

(1) Gas kinetic formulae for collision frequency cannot be used in the calculation of collision efficiencies in these processes. The discrepancy between the active life of the excited sensitizer as calculated by the application of simple gas-kinetic formulae to sensitization data and that calculated from intensity of absorption can be explained in principle in terms of the "cage" theory of collisions between solute molecules in liquid solutions, and may be adduced in support of this hypothesis.

(2) Steric requirements at collision might be expected to reduce efficiency, although in the reaction studied here this factor is unimportant.

(3) Fluorescence and sensitization, although always reciprocally related, may not be simple alternatives in the polyatomic case. Here avenues are open for the distribution of the excitation energy in states in the upper electronic level which cannot combine radiatively with the normal and therefore do not cause fluorescence, but which are still capable of transferring energy at collision, and therefore of sensitizing.

In the course of the argument, the fluorescence spectrum of naphthalene has been described with respect to the quantum yield as a function of exciting wave length in solution and in the gaseous state; the photochemical implications of an undiminished yield in a "predissociation" region of absorption as compared with that in the sharp absorption region for the solution, contrasted with its fall in the vapor, are discussed; data are given on the quenching of the fluorescence of naphthalene and certain derivatives by ethyl iodide, and on the quantum yield of the sensitized decomposition of ethyl iodide by naphthalene in hexane solution in its dependence on the concentrations of the reactants, the wave length of the exciting light and, qualitatively, on the viscosity of the medium.

DISCUSSION

The following discussion pertaining to material in this paper was contributed by members of the conference whose names and remarks appear below.

E. RABINOWITCH:[*]—West's observation that the actual quenching efficiency is substantially lower that that calculated from the collision frequency does not necessarily point to a small "steric factor." The discrepancy may be due to the unwarranted use of gas kinetic calculation methods. As shown elsewhere,[1] the notion of "collision frequency" should be supplemented, in dealing with condensed systems, by the notion of the "frequency of co-ordinations," or, to use the more convenient term suggested by Fowler and Slater,[2] "frequency of encounters." The reason is that two dissolved molecules, once they have met in an "encounter," remain "caged" together for the duration of several collisions,—until they finally get separated by diffusion. In other words, collisions occur in "sets." In treating reactions involving a considerable energy of activation (and therefore occurring only in one collision in a million or more), the non-uniform distribution of collisions (*i. e.* occurrence of collision sets, interrupted by long intervals) is irrelevant. For reactions which take place by the very first collision, the distinction between collisions and encounters acquires considerable importance. The frequency which determines the velocity of these reactions is the frequency of *encounters,* not that of the collisions. The frequency of encounters is equal to collision frequency divided by the average length of the collision sets; it is related to *diffusion velocity* and dependent on *viscosity* (whereas the collision frequency is, to the first approximation, independent of viscosity). The average length of collision sets in ordinary liquids is probably of the order of 10-100. It can thus well account for the major part of the discrepancy between collision frequency and quenching efficiency as described by West. The influence of viscosity, also described in West's paper, is an argument in favor of this explanation.

The assumption is often made that only fluorescent substances are capable of acting as sensitizers. This view is particularly widely held by biologists working on so-called "photodynamic action." Often such a correlation is undoubtedly borne out by facts, but it should not be construed as a general rule. The photosensitizing action of *ferric ions* in certain reactions, mentioned in today's discussion, is an illustration of the fact that non-fluorescent colored substances may also serve as sensitizers.

If we assume that *fluorescence* and energy *dissipation* (by internal conversion into vibrational energy, with ultimate transfer to the

---

solvent) are two competing *monomolecular* reactions, which together account for the de-activation of excited molecules in solution, the total velocity of de-activation becomes:

$$(1) \qquad\qquad K = K_f + k_d$$

where K_f is the velocity of de-activation by fluorescence and K_d that of de-activation by dissipation. The fluorescence yield is

$$(2) \qquad\qquad \varphi = \frac{K_f}{K_f + K_d} = \frac{1}{1 + K_d/K_f}.$$

If, now, a "sensitization substrate" is added, capable of interacting with the excited molecule in a *bimolecular* reaction, with a velocity constant K_s, the quantum yield of sensitization is

$$(3) \qquad\qquad \gamma = \frac{K_s[S]}{K_d + K_f + K_s[S]}.$$

For all typical dyes, *i. e.* substances with intense absorption bands in the visible region, K_f is of a similar order of magnitude (determined by the order of magnitude of the extinction coefficient), namely 10^7 to 10^8 sec.$^{-1}$. Differences in the fluorescence yield are therefore primarily determined by different velocities of energy dissipation K_d. This must be particularly true for closely related substances, like chlorophyll and Cu-chlorophyll, which have very similar spectra, but differ strongly in the intensity of their fluorescence. With a constant value of K_f, a low value of K_d is obviously favorable both for a high fluorescence yield φ (2) and a high sensitization yield γ (3). However, equation (3) shows that a considerable yield of sensitization may be obtained also when the yield of fluorescence is small, *i. e.* when $K_d \gg K_f$.

The condition for effective sensitization is, in fact,

$$(4) \qquad\qquad K_s[S] \gg K_d + K_f = \frac{K_f}{\varphi}.$$

This condition can be satisfied, even if φ is small, either by [S] being sufficiently large, or by K_f being sufficiently small. For dyes with the normal high value of K_f, and a low fluorescence yield φ, a sensitizing effect can be obtained only by using a very high concentration of the substrate. Non-fluorescent colored substances with comparatively weak extinction bands, and correspondingly low values of K_f ($\sim 10^6$ sec^{-1}) may, however, act as sensitizers even if the concentration of the substrate is not very high. The following table

TABLE 1

	Inorganic Ion					Typical Dye				
K_f			10^5					10^7		
K_d	10^{12}	10^{11}	10^{10}	10^9	10^8	10^{12}	10^{11}	10^{10}	10^9	10^8
φ	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}
$S_{1/2}$	10	1	0.1	0.01	0.001	10	1	0.1	0.01	0.001

gives an idea of the order of magnitude of the constants involved. We assume $K_s = 10^{11}$, which is a fair estimate for the velocity constant of reactions occurring by the very first encounter of sensitizer and substrate, and choose two values of K_f,—one typical of a weakly colored substance (maximum molar extinction coefficient of the order of 10^2–10^3) and one of a dye (maximum extinction coefficient of the order of 10^4–10^5). We give to K_d values from 10^{12} (dissipation of energy in a period of 1–10 molecular vibrations) to 10^8 (dissipation after 10^4–10^5 vibrations). This corresponds to fluorescence yields from 10^{-7} to 10^{-3} for the weakly colored substance, and 10^{-5} to 10^{-1} for the dye. The last row shows substrate concentrations which are required to obtain a sensitization yield $\gamma = 0.5$, according to (3). We recognize that sensitization with a weakly colored sensitizer may become efficient as $[S] = 0.001$ mol/l. even if the fluorescence yield is only 0.1%; whereas with a typical dye, the same efficiency at this concentration is correlated with a fluorescence yield of 10%. Concentration as high as 1 mol/l. is required to obtain efficient sensitization by a dye with a fluorescence yield of 0.01%.

Of course, a high concentration $[S]$ becomes unnecessary if the substrate forms a complex with the sensitizer. In photosynthesis, e. g., the fluorescence of the sensitizer (chlorophyll) is reduced to $< 10^{-2}$, whereas the sensitization efficiency is high. According to the table, the substrate concentration required under these conditions is 0.01 mol/l.; it is more probable, that the substrate forms a complex with chlorophyll, so that no encounters between sensitizer and substrate are required for the sensitization to take place.

Our consideration, of course, does not take into account cases in which the sensitizer is chemically changed (e. g. dissociated) by absorption, and the sensitization is caused by the reaction products (e. g. by iodine atoms produced by dissociation of I_2^*).

Several years ago, together with Dr. Beutler, we made some observations on the mercury-sensitized fluorescence of ammonia which we have had no occasion to publish. Dickinson and Mitchell[3] previously

³ Dickinson, R. G., & Mitchell, A. C. G. Proc. Nat. Acad. Wash. 12: 692. 1926.

have described this fluorescence as consisting of one band in the near ultraviolet, at about 3370 Å, and one band in the visible (with a maximum at about 5100 Å. We found no trace of the visible band and suspected that Dickinson's and Mitchell's observation might have been due to a fluorescence of the quartz walls. The band at 3370 Å, on the other hand, belongs to NH_3. It shows a very peculiar behavior in respect to pressure. With increasing NH_3-pressure, its diffuse boundary on the short wave side remains unchanged, but its rather sharp long-wave edge moves steadily towards the visible; at about $p = 20$ millimeters the band spreads into the visible region, and the Hg-NH_3 mixture begins to fluoresce with a *bluish*-white color, (very different from the *greenish* fluorescence of quartz).

The spread of fluorescence into the visible is not due to a simple increase in the general *intensity* of the fluorescent band, but to a one-sided shift whose explanation offers an interesting problem.

O. K. RICE: *—It is well known that in solution collisions between solute molecules occur in groups, one collision usually being immediately followed by many further collisions between the same pair, whereas in the gas phase each collision is a separate occurrence, essentially unrelated to past or future collisions. The total number of collisions can be calculated, in either solution or gas phase, by the "equilibrium" method. From this, the time to the next collision can be calculated directly for the gas phase but in solution a totally different type of calculation will be necessary to find the time before the next group of collisions, *i. e.*, the time until the *first* collision of the *group* occurs. This will depend on such factors as the viscosity of the solvent.

If any reaction requires an activation energy, then the probability of reacting at any collision is small. Hence the rate is determined by the total number of collisions, and can be calculated by an equilibrium method, as exemplified by the Brønsted equation, which gives the dependence on ionic strength in solution. If there is no activation energy required, then the determining factor in solution is the time until the first collision of a group of collisions. It remains to be shown whether the indicated calculation gives a result which depends in the same way on the ionic strength. Unless it does, or unless the quenching of fluorescence involves an activation energy, the interpretation of Stoughton and Rollefson's[4] experiments must remain in doubt.

* University of North Carolina.
[4] Stoughton, R. W., & Rollefson, G. K. Jour. Am. Chem Soc. 61: 2634. 1939.

E. TELLER:*—*Remark on the fluorescence of resonating ionic molecules.* Quenching of the fluorescence of resonating ionic molecules by electrolytes present in the solution has been discussed in terms of collisions between the resonating molecules and the ions of the electrolyte. Let us write the resonating molecules in the form $A \ R \ A^+$ together with the form $A^+ \ R^* \ A$ where R and R^* are identical radicals as far as the nuclear configuration is concerned but are mirror images with regard to their valence pictures. The electrostatic interaction between the charged atom which we have called A^+ and the ions in solution influences strongly the chance of the approach of an ion of the solution and subsequent quenching of the fluorescence. In calculating this effect due to Coulomb forces it is usually assumed that the ion A^+ carries a full electronic charge. The purpose of this remark is to show that this type of calculation neglects two important effects.

First in the resonating ion the charge is found only with the probability $\frac{1}{2}$ on one A-atom. Thus the Coulomb force is only half as large as has been assumed hitherto.

The second effect is concerned with the change of the charge distribution in the resonating ion caused by the approach of the quenching ion. The fundamental state of the resonating ion can be represented by the formula $\frac{1}{\sqrt{2}} (\psi_1 + \psi_2)$ where ψ_1 represents the electronic distribution in the $A \ R \ A^+$ configuration and ψ_2 stands for the proper function of $A^+ \ R^* \ A$. The excited state which is responsible for the color of the resonating ion is then $\frac{1}{\sqrt{2}} (\psi_1 - \psi_2)$. If a negative ion approaches the second A atom then the configuration $A \ R \ A^+$ will be represented more strongly in the ground state and ψ_1 will have a greater weight so that the corresponding proper function will become $\frac{1}{\sqrt{2}} [(1 + \epsilon) \psi_1 + (1 - \epsilon) \psi_2]$ where the number ϵ (which we suppose to be small for the sake of simplicity) is a measure of the polarizing influence of the approaching ion. Since the excited state must remain orthogonal to the ground state we obtain $\frac{1}{\sqrt{2}} [(1 - \epsilon) \psi_1 - (1 - \epsilon) \psi]$. Thus in the excited state ψ_2 and with it the configuration $A^+ \ R^* \ A$ has the greater weight though a negative ion has approached the second atom. It is seen that while for the funda-

* George Washington University.

mental state the polarization of the resonating ion is in the expected direction, the excited state is polarized in the opposite sense. For the quenching process interaction with the excited state must be considered and it follows that in this state an approaching negative ion causes the positive charge of the resonating ion to concentrate on the far A-atom. In this way the Coulomb attraction is reduced. Similarly one can see that if a positive ion approaches the excited resonating ion the positive change tends to concentrate on the adjacent A-atom and the Coulomb repulsion is increased.

This paradoxical effect can be expected only if the two discussed states of the resonating ion are close to each other and sufficiently removed from higher excited states. The influence of these higher states is always to produce polarization in the expected direction and as a rule this influence is considerable. But the tendency of a "wrong" polarizability or at least a diminished "right" polarizability has to be expected for low-lying excited states.

Influence of symmetry on the fluorescence of dyes. One way by which fluorescence can be inhibited is a process which may be called internal quenching. This process amounts to a conversion of the electronic excitation energy into vibrations (*i. e.* into heat). But that can happen as a general rule only if the potential surface of the excited state crosses the potential surface of the ground state[5] and if the molecular vibrations caused by the electronic transition can carry the molecules to the crossing point. In symmetrical molecules only totally symmetrical normal vibrations are strongly excited by the electronic transition. This restriction may make internal quenching impossible. Thus symmetrical molecules could be expected to show a greater tendency toward fluorescence.

Consider, for instance, the symmetrical ion resonating between the two forms $A^+ B A$ and $A B^* A^+$ (see "Remark on the fluorescence of resonating ionic molecules"). This ion will have a greater tendency to fluoresce than an asymmetric ion, for instance a similar ion in which one of the A-atoms has been replaced by a different atom. The attachment of a foreign ion to one of the A-atoms may cause a sufficient asymmetry to make internal quenching possible. This may perhaps explain in some cases the quenching of fluorescence of dyes by ions.

E. W. R. STEACIE:[*]—It should be noted that from the point of view of photosensitization experiments the occurrence of HgH as an inter-

[5] Teller, E. Jour. Phys. Chem. 41: 109. 1937.
[*] The National Research Council of Canada.

mediate in the dissociation of a hydrocarbon (or of hydrogen) is only of importance provided that the molecule formed has a fairly long life. If this is the case the HgH molecule can be formed in the normal state and receive the activation energy for its dissociation by collision. Under these conditions the reaction
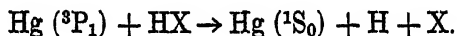
$$Hg + HX \rightarrow HgH + X$$
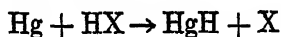
will be 8.5 kcal. less endothermic than

$$Hg + HX \rightarrow Hg + H + X.$$
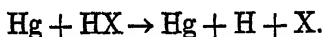
If, however, HgH is formed as an intermediate but dissociated before its next collision, *i. e.*

$$Hg\ (^3P_1) + HX \rightarrow HgH^* + X \rightarrow Hg\ (^1S_0) + H + X,$$
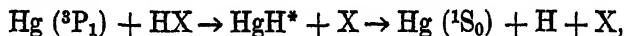
then the process is indistinguishable by photosensitization experiments from

$$Hg\ (^3P_1) + HX \rightarrow Hg\ (^1S_0) + H + X.$$

In any case, the spectroscopic evidence is not entirely clear cut, and I think that in a general way photosensitization experiments support the idea that the primary process always involves the sensitizer chemically and is never a mere transfer of energy. On this basis I prefer to consider HgH as an intermediate in the reaction of Hg ($^3P_1$) with both hydrogen and ethylene. This would also indicate why ethane dissociates by a C-H rather than a C-C bond split, since

$$Hg\ (^3P_1) + C_2H_6 \rightarrow HgH + C_2H_5$$
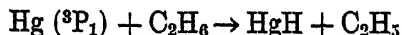
would seem more likely than

$$Hg\ (^3P_1) + C_2H_6 \rightarrow HgCH_3 + CH_3.$$
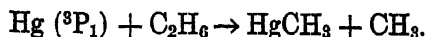
PHOTOLYSIS OF METAL ALKYLS AND THEIR SIGNIFICANCE IN PHOTOPROCESSES

By Hugh S. Taylor

From Princeton University

It is evident that whereas the data on the spectroscopy of the diatomic molecule have furnished much significant information with reference to the primary process in photochemical reactions, the corresponding data on the spectra of polyatomic molecules are by no means as compelling in this same respect.[1] As a consequence, recourse is had to the less direct and deductive processes that are possible when a photochemical process has been subjected to detailed and quantitative study. The present paper applies such a procedure to the processes occurring when metal alkyls undergo photolysis, either alone or in the presence of other reactants which do not participate in the primary process but which may react with the products of the photoprocess resulting from light absorption by the metal alkyl.

The primary process in the photolysis of metal alkyls has been shown in several investigations to be a convenient source of free alkyl radicals.[2, 3] Mercury diamethyl shows a continuum beginning at 2800 Å and extending into the far ultraviolet, upon which is superimposed a group of diffuse bands in the region of 2100-1970 Å. Terenin identified the alkyl by the metal mirror technique with light at 2100Å. Leighton and Mortensen[4] ascribed the reactions upon photolysis of lead tetramethyl to radical formation. Mercury dimethyl and diethyl as well as zinc diethyl have been used in a series of researches[5, 6, 7] to examine the reactions of the corresponding radicals in a variety of gaseous media unaffected by the light employed for photolysis of the alkyls. These studies have served still further to confirm free radical formation in the primary process of photolysis, to indicate that it is probably the exclusive process and to supply quantitative data on the reactivities of such radicals. The especial advantage of the metal alkyls is that their photo-decomposition

[1] See, for example, E. Teller's article in this symposium.
[2] Terenin, A. Jour. Chem. Phys. 2: 441. 1934; Acta Physico-chim. USSR 1: 762. 1935 Terenin, A., & Prileshajeva, N. Trans. Faraday Soc. 31: 1483. 1935.
[3] Linnett, J. W., & Thompson, H. W. Trans. Faraday Soc. 33: 501, 874. 1937.
[4] Leighton, P. A., & Mortensen, R. A. Jour. Am. Chem. Soc 58: 448 1936.
[5] Cunningham, J. P., & Taylor, H. S. Jour. Chem. Phys 6: 359. 1938.
[6] Smith, J. O., & Taylor, H. S. Jour. Chem. Phys. 7: 390. 1939. Jour. Chem. Phys 8: 543 1940.
[7] Moore, W. J., & Taylor, H. S. Jour. Chem. Phys. 8: 396. 1940.

produces only hydrocarbons or their fragments and the metallic atoms, which latter do not normally complicate the reaction system.

Mercury dimethyl, between the temperatures of 40 and 200°, produces on photolysis practically exclusively ethane.[5] That this involves a free radical mechanism and not a direct intramolecular formation of the ethane is evident from the behavior in presence of hydrogen. In this case methane is formed at all temperatures and in increasing amounts with increase of temperature. The activation energy of the reaction is 9 ± 2 Kcals. In the interval of temperature from 250 to 300° the methane formed accounts for more than two thirds of the total hydrocarbon product when hydrogen is present in a five-fold excess.

With mercury diethyl,[7] the hydrocarbon products are more complex. In addition to butane, ethane and ethylene appear in approximately equivalent amounts at room temperatures and with ethane appearing in excess as the temperature rises. Some butylene and minor amounts of hydrogen are also formed. Again, however, the free radical mechanism of the primary photolysis is evident from the reaction in presence of hydrogen. The process is similar to that with methyl radicals and has a similar activation energy. At 250° the increasing hydrogen pressure causes a major increase in ethane formation, as shown in the following data:

TABLE 1

ETHANE YIELD FROM PHOTOLYSIS OF MERCURY DIETHYL WITH VARYING HYDROGEN PRESSURES

| $T = 250°$ C.; P (Hg $(C_2H_5)_2$) $= 25$ mm.				
P_{H_2} 0	50	100	200	700
C_2H_6 % 39	48	56	62	90
C_2H_4 20	14	12	8	1
$C_4H_8 + C_4H_{10}$ 41	38	32	30	9

The ethane yield at the highest hydrogen pressure corresponds to more than 80% conversion of the ethyl radicals in mercury diethyl to ethane and is a direct proof that a major, if not exclusive, product of the photolysis is the free radical.

Independent confirmation of a primary formation of radicals is to be had from the data on quantum yield. Linnett and Thompson[3] found for the photodecomposition of mercury dimethyl a quantum yield of approximately unity at room temperatures increasing to 2.6 around 200° C. A similar threefold increase in a 200° temperature

range was found by Moore and Taylor[7] for the photodecomposition of mercury diethyl. When these alkyls are decomposed in the presence of hydrogen there is a further twofold increase in the velocity of photodecomposition[5, 7] pointing to quantum yields of as much as 5 in the region of 200–250° C. Since the possibility of a decomposition to yield mono-alkyl as the primary process cannot be excluded

$$R_2Hg + h\nu = RHg + R$$
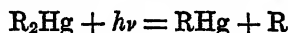
and consequently, also, the possibility of the reverse process of radical recombination,

$$R + RHg = R_2Hg,$$
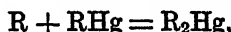
it is evident that a quantum yield of unity at room temperatures does not exclude a chain mechanism and the higher quantum yields at elevated temperatures certainly indicate a chain reaction. A primary process to yield saturated hydrocarbon plus metal atom

$$R_2Hg + h\nu = Hg + R_2$$
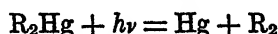
or a reaction with hydrogen to yield two molecules of a saturated hydrocarbon

$$R_2Hg + h\nu = (R_2Hg)^1$$

$$(R_2Hg)^1 + H_2 = 2RH + Hg.$$
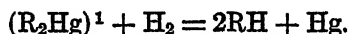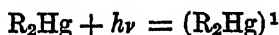
are not compatible with a chain mechanism or quantum yields exceeding unity.

On the assumption that the primary process is an exclusive formation of radicals the reactivity of methyl radicals towards a variety of hydrocarbons has been examined by measurement of the extent of formation of methane[6] which is a very minor product below 200° in the photolysis of mercury dimethyl. Over a range of temperatures from 100 to 300° C. the activation energy of methane formation by interaction with the hydrocarbons has been determined with the following results.

TABLE 2

ACTIVATION ENERGIES OF METHANE FORMATION FROM METHYL RADICALS AND HYDROCARBONS

Hydrocarbon	E Kcal.	Hydrocarbon	E Kcal.
Ethane	8.1	Ethylene	>10
Neopentane	8.3	Propylene	3
Butane	5.5	Benzene	>10
Isobutane	4.2	Toluene	5.6
		Diphenyl methane	< 5.6

It is of interest that the variation of these activation energies with the molecular structure of the hydrocarbon parallels entirely the results obtained by Polanyi and his collaborators in their studies of the interaction of sodium vapor with a variety of organic halides. There is a definite increase in reactivity as one passes from primary, through secondary to tertiary carbon-hydrogen (or carbon-halogen in Polanyi's work). In both sets of experiments decrease in activation energy is found for compounds with a carbon double bond in a position once removed from the atom linkage attacked and an increase in activation energy or retardation of reaction rate if the reacting atom is directly attached to the doubly-bonded carbon atom. This applies not only to the olefines but also to the aromatic series (*cf.* benzene and toluene).

The differences in activation energy between ethylene and propylene on the one hand and between benzene and toluene on the other may be associated with the resonance characteristics of the allyl and benzyl radicals produced, from propylene and toluene, respectively, by withdrawal of a hydrogen atom. It has been estimated[8] that the resonance energy of the allyl radical amounts to 15.4 Kcal. so that this energy would separate the potential energy surface of the radical-methane product from that which would obtain with a non-resonating radical-methane product from another hydrocarbon reaction as, for example, that between methyl and ethylene. A similar energy difference, due to resonance, of the activated complexes in the two reactions would account fully for the differences in methane formation observed. Similar differences between the well-known resonating structure, the benzyl radical, and that of phenyl would account for the differences in reactivity between toluene and benzene. For the same reason, the reactivity of methyl radicals towards diphenyl methane ought to be greater than that towards toluene, as actually found.

The interpretation of the experimental data with respect to the unsaturated and aromatic hydrocarbons on the basis of resonance energy of the activated complexes and the radicals produced rather than on the basis of bond strengths of the compounds studied raises the question as to whether the data on the reactivities of the primary, secondary and tertiary C-H groupings in the saturated hydrocarbons might not be similarly interpreted. In these cases the radicals formed, *e. g.*, ethyl, butyl, isobutyl, are not ordinarily regarded as resonating struc-

8 Lennard-Jones, J. E., & Coulson, C. A. Trans. Faraday Soc. 35: 811. 1939.

tures. The differences in activation energies observed, amounting to 4 Kcal. between ethane and isobutane are of a lower order of magnitude than those observed between ethylene and propylene. By attributing these differences of activation energy to differences of stability due to resonance effects between such radicals as ethyl and isobutyl, one could retain the Pauling principle of C-H bond energies practically identical in magnitude in the saturated hydrocarbons. The thermochemical data of Kistiakowsky and his co-workers do however suggest that differences in bond energies up to but not exceeding those found in our earlier work may exist in such systems.

A more difficult problem is involved, however, in the application of resonance to the comparative rates of formation of methane from the interaction of methyls with ethane and ethylene. The greater reactivity of ethane would suggest that there was a much greater resonance in the activated complex $[CH_3-C_2H_6]$ than in that of $[CH_3-C_2H_4]$. It is not at once evident why this should be so. The alternative is that the C-H bond strength in ethylene is stronger than that in ethane. This conclusion is in accord with classical views in organic chemistry. It is deriving support from the work now in active prosecution which interprets the restricted rotation in a molecule such as ethane in terms of the effect of the methyl radical on the adjacent C-H bonds, tending to weaken these. The same processes of reasoning would lead to a strengthening of the C-H bonds in ethylene. As already emphasized the problem is not peculiar to C-H systems but extends, as in the work of Polanyi, to the carbon-halogen linkages in various homologous groups of organic compounds.

It is further evident that these variations in reactivity of olefines and aromatic hydrocarbons with free radicals indicate varying efficiencies of such materials as inhibitors of chain processes of decomposition. It is not surprising that different workers have found markedly different inhibitory efficiencies of ethylene and propylene on thermal decomposition processes. Even small differences in activation energies of competing reactions may cause large ratios of reaction due to the influence of the exponential term, $\exp (\Delta E/RT)$.

A further problem in this general field is the variation in products from a variety of photolytic and thermal processes yielding free ethyl radicals. A table compiled by Moore[7] is reproduced below and shows that the products may vary from 100% butane to 100% ethane + ethylene.

The best evidence available at the present time indicates that collision of two ethyl radicals produces butane. The disproportionation

TABLE 3

The Products from Ethyl Radicals Prepared from Various Sources

| Source of Ethyls | Temp. | Moles per cent in Product of | | | | | |
		H_2	CH_4	C_2H_4	C_2H_6	C_4H_{10}	C_4H_8
EtBr + Na, EtCl + Na (Bawn and Dunning)	350°	0	0	0	0	100	0
Et$_2$ CO photolysis (Noyes and Ells)	30	0	0	10	10	80	0
Mercury photosensitized hydrogenation of ethylene	25	—	4	—	16	80	0
Mercury diethyl photolysis	45	2	0	25	29	38	7
Zinc diethyl photolysis	45	10	0	25	22	34	9
Lead tetraethyl pyrolysis (Meinert)	150	1	0	38	46	14	2
Ethyl iodide + Hg vapor photolysis	100	1	0	29	43	15	12
Ethyl iodide photolysis (West)	25	0	0	50	50	0	0

reaction, $2C_2H_5 = C_2H_4 + C_2H_6$ would appear to have an activation energy at least 4 Kcal. greater than radical recombination and will attain significance in competition with the recombination only at the higher temperatures prevailing in the pyrolyses of hydrocarbons. The formation of ethane and ethylene observed in the photolysis of ethyl iodide, the considerable formation from the metal alkyls and the small quantity from diethyl ketone must, at present, be ascribed largely to reactions of the type

$$C_2H_5 + (C_2H_5)_nX = C_2H_6 + C_2H_4 + (C_2H_5)_{n-1} X. \qquad (A)$$
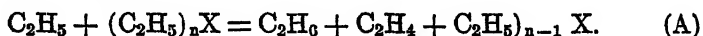
The activation energy for this reaction will be at most 2 or 3 Kcals. greater than that for the combination of ethyls to form butane.

A large variation in the ethyl radical concentration from the various sources of production is probably an important factor responsible for the large variation in products (*cf.* ethyl iodide and diethyl ketone). Because of the greater probability of formation of the activated state[9] an ethyl radical will combine with an iodine atom much more efficiently than with another ethyl or with a propionyl radical. Reaction (A) will therefore be favored in the case of ethyl iodide. Similar considerations as to the probability of the activated state indicate that the reaction of methyl radicals with hydrogen molecules is indeed the reaction $CH_3 + H_2 = CH_4 + H$ and not, as suggested by

[9] Bawn, C. E. H. Trans. Faraday Soc. 31: 1536. 1935.

H. A. Taylor and Burton[10] a reaction, $2CH_3 + H_2 = 2CH_4$, without formation of atomic hydrogen. Such a mechanism as this latter is also not compatible with the increased chain length when metal alkyls are photolysed in the presence of hydrogen, since the reaction $2CH_3 + H_2 = 2CH_4$ cannot set up a chain sequence. As has been shown recently by Taylor and Smith,[6b] when mercury dimethyl is photolysed in presence of nitrogen and of hydrogen, the ethane yield is the same with the two gases. With hydrogen, however, the yield of methane is twelve-fold that in presence of nitrogen. The increased methane yield must represent the influence of the operation of the chain mechanism.

In TABLE 4 are recorded the products obtained in the photolysis of methyl ethyl ketone in the temperature range of 90–200° C.[11] The lower limit of temperature was set by the desire to avoid the complicating formations of diketones.

TABLE 4

PERCENTAGE COMPOSITION OF PHOTOLYSIS PRODUCTS OF MeEtCO

Temp. °C.	90	110	155	200
H_2	0.4	0.1	0.3	0.0
CO	47.5	49.0	47.0	48.5
CH_4	3.4	4.4	9.0	18.0
C_2H_4	2.2	0.5	1.7	3.0
C_2H_6	11.5	13.0	14.5	12.5
C_3H_8	19.1	20.0	18.6	11.0
C_4H_{10}	13.2	13.0	9.0	7.0
C_4H_8	1.5	—	0.0	—

It will be seen that the formation of ethane, propane and butane parallels the frequency of the corresponding free radical collisions. This evidence suggests a primary photolysis to yield free radicals. At the higher temperatures, production of ethane and especially of methane at the expense of higher hydrocarbons becomes predominant, due doubtless to secondary reactions between radicals formed by photolysis and undecomposed ketone.

DISCUSSION

The following discussion pertaining to material in this paper was contributed by members of the conference whose names and remarks appear below.

[10] Taylor, H. A., & Burton, M. Jour. Chem. Phys. 7: 675. 1939.
[11] Moore, W. J., & Taylor, H. S. Jour. Chem. Phys. 8: 466. 1940.

F. O. RICE:*—It seems likely that metal atoms such as mercury or zinc would play a part in the radical reactions that follow the primary production of radicals from the organo-metallic compound. The structure of these atoms is such that one might expect a readiness to react with an alkyl radical. In any case it is very probable that it is the lead atoms that so strongly affect the course of the reaction in the oxidation of hydrocarbons. Possibly one might use some of the higher lithium alkyls as the source of radicals.

W. WEST:*—Evidence that the energy of activation of the reaction $CH_3 + H_2 = CH_4 + H$ is comparatively low, say below 10,000 cal. raises doubt as to the mechanism of the ortho-para hydrogen conversion in the illumination of ortho-para mixtures with methyl iodide.[1] Both H atoms and free radicals can cause conversion, the former with an intrinsic efficiency of about a million times that of the radicals. Certain experiments of Paneth[2] on the half life of methyl radicals seemed to show the same half life in hydrogen as transporting gas as in helium or nitrogen, suggesting the improbability of reaction between methyl and molecular hydrogen, and the conversion observed in illuminated methyl iodide was therefore attributed to the paramagnetic properties of radicals arising from photo-dissociation. But streaming experiments such as those of Paneth are probably inadequate to exclude the possibility of reaction between methyl and hydrogen with some energy of activation. Indeed the figures given in the paper by Paneth and Herzfeld suggest a shorter life in hydrogen than in inert gases, though the authors explicitly refrain from regarding the differences as significant. An energy of activation of 8000 cal. would correspond to one collision in a million between methyl and molecular hydrogen leading to reaction, and if the life-periods of H and CH_3 were about the same, about half the observed conversion in the photo-experiments would be due to the atoms. In view of the evidence that the life period of hydrogen atoms in hydrogen is many times that of methyl radicals, it is necessary to conclude that if the reaction between methyl and molecular hydrogen has an activation energy below about 10,000 cal., the bulk of the conversion of ortho to para hydrogen in reactions involving the production of methyl will be due to hydrogen atoms. The data at

[1] West, W. Jour. Am. Chem. Soc. 57: 1931. 1935

[2] Paneth, F., & Herzfeld, K. Zeit. Elektrochem. 57: 378. 1931.

present available on this conversion by illuminated methyl iodide are not sufficient to allow a satisfactory estimate to be made of the relative amounts of conversion due to hydrogen atoms and radicals, though such experiments made with careful measurement of the rate of absorption of light would be very significant in this connection. It need scarcely be pointed out that however much the mechanism of the conversion is of interest in itself, its details do not affect the validity of the conversion as a means of detecting the production of free radicals so long as hydrogen atoms are not produced in other ways than by the action of the radicals. In fact the effect of the occurrence of the reaction $CH_3 + H_2 = CH_4 + H$ is to enhance the sensitivity of the conversion in the detection of methyl.

E. W. R. STEACIE:[*]—The best estimate of the activation energy of the exchange of D atoms with methane is probably about 13 Kcal. However, there are two possible mechanisms for the exchange,

$$D + CH_4 \rightarrow CH_3D + H \tag{1}$$

and

$$D + CH_4 \rightarrow CH_3 + HD. \tag{2}$$
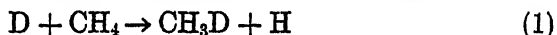
If, as some experiments seem to indicate, (1) is the mechanism of the exchange, then it is (1) which has an activation energy of ca. 13 Kcal., and the activation energy of (2) may be considerably greater than this value. However, it seems very unlikely that (1) is the true mechanism, and the evidence strongly favors (2).

H. AUSTIN TAYLOR:[*]—In the reaction between methyl radicals and ethylene it seems to me that the rates to be compared are those of the alternative rearrangements of the activated complex to give on the one hand propyl radicals and on the other methane and C_2H_3.

[I] The National Research Council of Canada.
[*] New York University, New York City, New York.

THE AMPHOTERIC PROPERTIES OF PROTEINS[*]

By

R. Keith Cannan, A. Kibrick, John G. Kirkwood, L. G. Longsworth,
A. H. Palmer, and Jacinto Steinhardt

CONTENTS

THE AMPHOTERIC PROPERTIES OF
EGG ALBUMIN

By R. Keith Cannan, A. Kibrick, and A. H. Palmer

From New York University, College of Medicine, New York City, New York

This paper is devoted to a discussion of the contribution which electrode titrations may make to the quantitative definition of the amphoteric properties of a protein. The argument will be illustrated by experimental curves for the titration of egg albumin with HCl and with NaOH under a variety of conditions. Some reference will also be made to a comparable study of β-lactoglobulin,[1] the detailed results of which will be published elsewhere.

The curves, which will be discussed, represent the relation between pH and a quantity which we will designate h. This quantity is simply the difference between the equivalents of HCl ($+$ ve) or of NaOH ($-$ ve) added to one mole of protein and the equivalents of free H^+ ($+$ ve) or of OH^- ($-$ ve) present at the observed pH. We have adopted a value of 45,000 for the molecular weight of egg albumin.[2]

The corrections for free H^+ and OH^- have been computed on the assumption that,

$$pH = - \text{Log} [H^+] \gamma_{H^+} = \text{Log} [OH^-] \gamma_{OH^-}/K_W$$

The activity coefficients used have been calculated from observations of the pH of dilute protein-free solutions of HCl and NaOH containing amounts of inorganic salts corresponding to those in the various protein systems. (Table 1.) This method involves the assumption that the presence of protein does not affect the magnitude of either the activity coefficients or the liquid junction potentials. This, however, can be a significant source of error in the value of h only at the extremities of the curves. Indeed, it is the rapidly mounting importance of the correction for free H^+ or OH^- at these extremities which fixes the effective limits of titration at about pH 2 and 12. Within these limits, we are satisfied that the pH of a given protein-HCl-NaOH-salt mixture can be reproduced on independent protein preparations with a probable error of less than $\pm$ 0.02 pH— except in the substantially unbuffered region between 7.5–9.5.

---

[1] β-lactoglobulin is the name which we propose for the crystalline globulin which Palmer (Jour. Biol. Chem. **104**: 359. 1934) isolated from the albumin fraction of cow's milk.

[2] Tiselius, A., & Svensson, H. Trans. Faraday Soc. **36**: 16. 1940.

The significance of h depends upon the assumptions which are made with respect to the nature of the reaction with acid and base. If the reaction is confined to H^+ and OH^-, then h is equal to the number of H^+ combined with ($+$ ve) or dissociated from ($-$ ve) one mole of initial protein. If the isoionic condition of the protein can be identified and h is measured from this condition, h, then, corresponds with the mean valence (z) of the protein. Should, however, the protein combine with ions other than those of water, $h = z + a - c$, where a and c are the respective equivalents of bound anions and cations. One further qualification must be made. The occurrence of any irreversible reaction of the protein with H^+ or OH^- will modify the significance of the curves with respect to the amphoteric properties of the original protein. Moreover, should such irreversible reactions attain significant rates in specific regions of the curve, the shape of those regions will depend upon the conditions of observation. We are satisfied that the curves for egg albumin and

TABLE 1

μ	Calc. L.J. Potential	$-\text{Log } \gamma_{H^+}$	$-\text{Log } \dfrac{K_W}{\gamma_{OH^-}}$
2.38 M	-0 3 mv.	-0 077	13 88
1.19 M	-0 6 mv.	$+0.060$	13.85
0.667 M	-0 9 mv.	0 115	13 84
0.267 M	-1.4 mv.	0.125	13 85
0.133 M	-1.8 mv.	0.113	13 87
0.067 M	-2.1 mv.	0 095	13 90
0.033 M	-2 5 mv.	0.083	13 92
0.0167 M	-2 9 mv.	0.072	13 94
0.0085 M	-3 2 mv.	0 060	13 96

for β-lactoglobulin represent reversible equilibria over the range of about pH 2–11. Beyond these limits there is evidence for the participation of irreversible reactions.

The identification of the isoionic point of β-lactoglobulin presents no difficulties. This protein may be crystallized from solutions which are practically free from non-protein ions. The pH of a solution of the crystals is consistently found to be 5.20 ± 0.01. This is practically identical with the electrophoretic isoelectric point of 5.19 in 0.1 M acetate.[3] It is identical also with the pH of intersection of

[3] Pedersen, K. O. Biochem. Jour. **30**: 948. 1936.

a family of dissociation curves of β-lactoglobulin corresponding to a range of concentrations of KCl varying from 0.01 to 0.67 M.

The situation with egg albumin is more ambiguous. Egg albumin crystallizes from appropriate salt solutions over a wide range of pH (4.0–7.5). A solution of the crystals, even after protracted dialysis, stubbornly retains a proportion of the anions or cations originally bound to the crystals. Presumably exhaustive electrodialysis should produce isoionic material, but this may not be a convenient method for preparations in bulk. Sorenson, Linderstrøm-Lang, and Lund[4] prepared egg albumin by dialysis against dilute ammonia until free from sulphate, followed by a quantitative neutralization of the ammonia present with HCl. The pH of the resulting solution was 4.88 at 18°. They identify this with the isoionic point. Significantly, they found that a group of dissociation curves corresponding to a wide range of concentrations of NH_4Cl (0.05–2.96 M) intersected at a point very close to 4.9. We have adopted 4.90 as the Isoionic Point of egg albumin, although we shall later point out that this point on the curve is not independent of [KCl]. More specifically, our assumption is that $h = 0$ in an egg albumin solution which has been dialyzed substantially free from sulphate and has then been adjusted to pH 4.9 by addition of HCl or of NaOH. The amount of acid or base required by each preparation has been determined and all calculations of h have been corrected accordingly.

METHODS AND RESULTS

Our discussion of the amphoteric properties of egg albumin will be based upon studies of the effects on the curve of varying (a) the concentration of protein, (b) the concentration of KCl or of MCl_2 present (M = Ca, Sr, Ba or Mg), and (c) the temperature. We will also use information from published and unpublished work on the effect of formaldehyde on the curve.[5]

The main part of the work has consisted of an extensive investigation, over the whole experimental pH range, of the effect of varying [KCl] from 0.0085 to 2.38 M. The observations on MCl_2 are less extensive and cover only the range pH 2.5–6. In this work on the effect of salts the hydrogen electrode was employed at 25°. The saturated calomel electrode was standardized with 0.01 M HCl in 0.09 M KCl, assuming for this solution a pH of 2.10 and a liquid

⁴ Sorensen, S. P. L., Linderstrøm-Lang, K., & Lund, E. Jour. Gen. Physiol. **8**: 543. 1925–7.

⁵ Kekwick, R. A., & Cannan, R. K. Biochem. Jour. **30**: 235. 1936.

junction potential (L. J. P.) at a saturated KCl junction of − 2.2 mv.[6] In an attempt to compensate for changes in L. J. P. with changes in the salt concentration of the protein systems, we have subtracted, from the observed electrode potentials, L. J. P.'s calculated with the aid of Henderson's equation, using 74.2 and 77.2 for the equivalent conductivities of K^+ and Cl^- respectively (TABLE 1).

TABLE 2 presents a summary of the results. It was compiled in the following way. All observations at each [KCl] were assembled as points on a h : pH curve and a free-hand curve was drawn through them. The values in TABLE 2 were obtained by interpolation in these curves. Each curve is based upon observations on 40 or more reaction mixtures independently prepared. Four different preparations of egg albumin have been used in the work.[7] The precision with which TABLE 2 reproduces the observations may be indicated by the statement that no reproducible observation is displaced from its curve by more than 0.5 eq. The great majority of the points falls within 0.2 eq. of the curve. In FIGURE 1, several of the curves are reproduced to illustrate the general nature of the salt effect.

The effect of protein concentration has been examined both in 0.267 M KCl and in solutions containing no added ions other than those of the titrating agents. In the first case a change of protein concentration from 0.4 to 5% was without observable influence. In the absence of KCl an effect of [protein] was clearly evident. Data for three curves are reproduced in TABLE 3 which was constructed in the manner of TABLE 2. The curve for 2.5% protein is the published curve of Kekwick and Cannan[7] and is based on hydrogen electrode studies. The observations upon which the two other curves are based were obtained with the aid of a glass electrode-vacuum tube assembly. From these curves we will later attempt to estimate the magnitude of the contribution of protein ions to the ionic strength.

A summary of the results of observations at 3 different temperatures is shown in FIGURE 2. In this work a glass electrode was also employed. It was calibrated in the manner described by Wyman.[8]

DISCUSSION

Discussion of these curves falls naturally into three parts: (a) a stoichiometric analysis, (b) an empirical analysis of salt effects, and (c) a theoretical approach to these empirical relations.

[6] Guggenheim, E. A., & Schindler, T. D. Jour. Phys. Chem. **38**: 533. 1934.
[7] Kekwick, R. A., & Cannan, R. K. Biochem. Jour. **30**: 227. 1936.
[8] Wyman, J., Jr. Jour. Biol. Chem. **127**:1. 1939.

TABLE 2

DISSOCIATION CURVES OF EGG ALBUMIN AT CONSTANT μ (KCl)

h = equi. [H$^+$] bound per 45,000 gm. Egg Albumin

h	pH								
	2.38μ	1.19μ	0.67μ	0.267μ	0.133μ	0.067μ	0.033μ	0.017μ	0.0085μ
39	2.53	2.43	2.40	2.03	—	—	—	—	—
38	2.80	2.68	2.58	2.31	2.17	2.00	—	—	—
36	3.04	2.95	2.88	2.63	2.48	2.28	2.00	—	—
34	3.24	3.16	3.09	2.83	2.70	2.50	2.24	—	—
32	3.40	3.31	3.24	3.02	2.88	2.70	2.46	—	—
30	3.54	3.45	3.38	3.18	3.04	2.86	2.65	—	—
28	3.66	3.57	3.50	3.32	3.18	3.03	2.81	—	—
26	3.77	3.68	3.61	3.43	3.32	3.17	2.96	—	—
24	3.87	3.79	3.72	3.55	3.45	3.31	3.11	—	—
22	3.97	3.89	3.82	3.67	3.57	3.44	3.26	—	—
20	4.07	3.99	3.92	3.78	3.69	3.57	3.41	3.29	3.13
15	4.32	4.23	4.17	4.05	3.98	3.88	3.75	3.66	3.55
10	4.56	4.48	4.42	4.33	4.28	4.21	4.12	4.06	3.96
8	4.66	4.58	4.52	4.44	4.40	4.34	4.26	4.22	4.14
6	4.76	4.69	4.63	4.56	4.53	4.48	4.41	4.38	4.32
4	4.87	4.79	4.75	4.69	4.65	4.62	4.57	4.55	4.50
2	4.99	4.92	4.87	4.83	4.79	4.77	4.74	4.72	4.69
0	5.13	5.05	5.00	4.97	4.95	4.93	4.92	4.91	4.90
− 2	5.29	5.21	5.17	5.15	5.14	5.12	5.14	5.14	5.14
− 4	5.47	5.39	5.35	5.35	5.33	5.33	5.36	5.38	5.39
− 6	5.70	5.62	5.58	5.57	5.56	5.59	5.62	5.65	5.68
− 8	5.98	5.90	5.86	5.87	5.89	5.91	5.96	6.00	6.05
−10	6.39	—	6.26	7.28	6.29	6.33	6.39	6.45	6.54
−12	6.93	—	6.80	6.82	6.87	6.91	6.97	6.90	7.12
−14	(8.0)	—	(7.8)	(7.9)	(7.9)	(8.0)	(8.1)	—	—
−16	(9.2)	—	(9.0)	(9.1)	(9.1)	(9.2)	(9.4)	—	—
−18	9.75	—	9.62	9.69	9.73	9.80	9.94	—	—
−20	—	—	9.93	10.02	10.06	10.15	10.29	—	—
−22	—	—	10.15	10.26	10.31	10.41	10.56	—	—
−24	—	—	10.23	10.47	10.53	10.63	10.79	—	—
−26	—	—	10.52	10.65	10.73	10.84	11.02	—	—
−28	—	—	10.69	10.83	10.92	11.04	11.22	—	—
−30	—	—	10.88	11.02	11.12	11.23	11.42	—	—
α_s	0.016	0.0155	0.014	0.009	0.005	0.000	−0.007	−0.014	−0.0205
β_0 (KCl)	0.22	0.14	0.10	0.07	0.04	0.03	0.02	0.01	0.00
β_0 (CaCl$_2$)	—	—	−0.14	−0.10	−0.08	−0.06	−0.05	—	—

| [Protein]	80.3 gm. p. litre = 1.8 mM				32 gm. p. litre = 0.71 mM				circa 25 gm. p litre = 0.56 mM			
h	pH	μ	μ_s	μ_p	pH	μ	μ_s	μ_p	pH	μ	μ_s	μ_p
30	2.81	.063	.028	.035	2.62	.032	.012	.019	2.54	.024	.010	.014
26	3.06	.047	.024	.023	2.86	.022	.010	.012	2.80	.017	.008	.009
22	3.26	.032	.020	.012	3.11	.016	.008	.008	3.04	.012	.007	.005
18	3.48	.022	.017	.005	3.36	.012	.007	.005	3.29	.008	.005	.003
14	3.72	.016	.013	.003	3.63	.008	.005	.003	3.57	.005	.004	.001
10	4.01	.012	.009	.003	3.92	.005	.004	.001	3.88	.003	.003	—
6	4.33	.009	.005	.004	4.26	.0025	.002	.0015	4.25	.003	.002	.001
2	4.69	.005	.002	.003	4.68	.0025	.001	.0015	4.67	.001	.001	—
0	4.90	—	—	—	4.90	—	—	—	4.90	—	—	—
−2	5.14	.005	.002	.003	5.14	.004	.001	.003	5.16	.001	.001	—
−6	5.66	.012	.006	.006	5.69	.007	.002	.005	5.72	.004	.002	.002
−10	6.40	.024	.009	.015	6.45	.014	.004	.010	6.44	.006	.003	.003
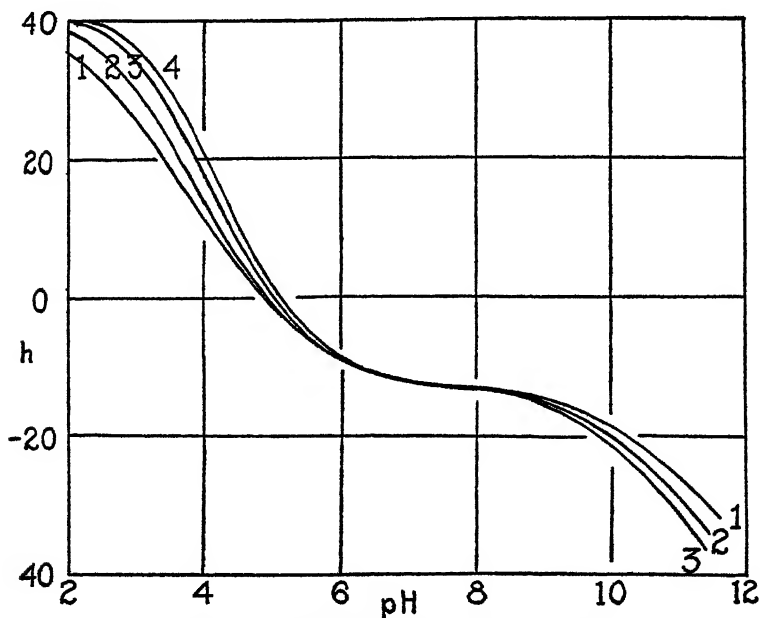
FIGURE 1.   Dissociation curves in varying [KCl].

Curve 1—0.033 M KCl.
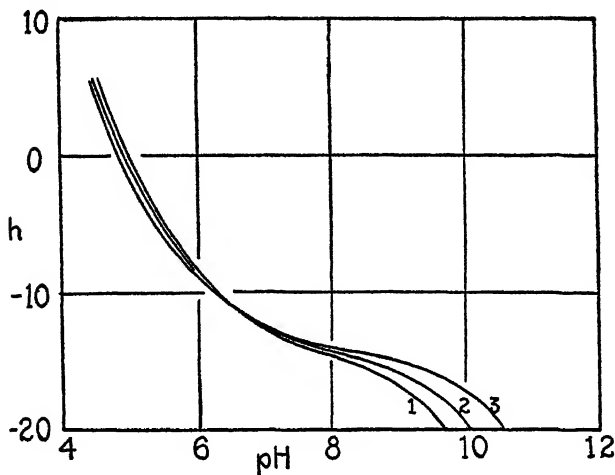Curve 2—0.133 M KCl.
Curve 3—0.667 M KCl.
Curve 4—2.38 M KCl.

Dissociation curves at varying temperature (KCl = 0.067 M).
Curve 1—43° C.
Curve 2—27° C.
Curve 3—7° C.

STOICHIOMETRY

On any plausible theory of the structure of egg albumin, the only ionizable groups whose presence is to be anticipated are the carboxyl, imidazole, amino, phenolic hydroxyl, and guanidine groups. We may expect the contributions of these groups to the dissociation curve to be centered in the pH ranges 3–5, 5–7, 9–11, 10–12, and 11–13, respectively. In the presence of formaldehyde the only important change in this distribution will be a shift of the amino groups from pH 9–11 to about pH 6–8. Three stoichiometric points in the curve should, therefore, be sought. One, at a pH of 2, or less, will correspond with the total cations; another above pH 13 with the total anions; and the third, in the region of pH 8.5, will represent a net charge equal to the total carboxyl groups less the cations present at this pH. In water the latter will be the sum of the amino and guanidine groups. In the presence of formaldehyde the guanidine groups will be the only cations at this pH.

The experimental curves for egg albumin are clearly approaching a common maximum at or below pH 2 (FIGURE 1). A special series of observations in this region on systems containing high concentrations of protein and of KCl, have led us to fix this maximum at $h = 41 \pm 0.5$. The published curves of Kekwick and Cannan,[7] for systems containing no KCl, give a value of 40 ± 1. At the alkaline

end, the feasible limit of accurate titration is at about pH 12. At this extreme there is no clear indication that the curves are approaching a limited value of h—a fact which is consistent with the presumption that the charge on the guanidine will not be completely suppressed below pH 13. Consequently, no direct measure of the total anions is possible. A well-defined break in the curves is, however, manifest at pH 8–8.5. In aqueous systems this corresponds with $h = -15 \pm 0.5$ while, in the presence of excess of formaldehyde, the value is -37 ± 0.5. The latter is a slightly higher value than that reported by Kekwick and Cannan[5] but is more dependable in that it was derived from observations on systems at constant ionic strength and
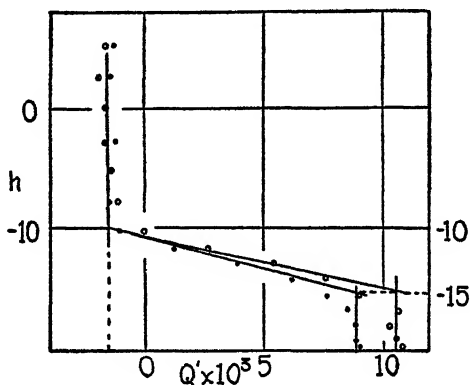

FIGURE 3.   Apparent heat of dissociation (Q').
● Temperature interval 7–27°.
○ Temperature interval 27–43°.

containing higher concentrations of formaldehyde than those used in the earlier work. The difference between the net charge in water and in formaldehyde at pH 9 is 22. It is a measure of the amino groups which react with formaldehyde.

Wyman[8] has shown how the contribution of the imidazole groups to the dissociation curve of a protein may be identified by determinations of the curve at different temperatures. Wyman assumes that the change in pH with temperature (ΔpH), at any chosen value of h, may be used to calculate an apparent heat of ionization (Q') from the approximation:

$$Q' = -4.79 \ \Delta pH \cdot \frac{T_1 T_2}{T_2 - T_1}.$$

He considers that Q' will reflect the type of group whose dissociation

dominates the curve in the region examined and he assigns values of about -1000, $+6000$ and $+10,000$ cals. to the COOH, imidazole and amino groups, respectively. The relation of Q' to h for egg albumin is shown in FIGURE 3. We conclude from this curve that the reaction of egg albumin with H^+ is predominantly due to COOH groups at values of h positive to -10 and predominantly due to amino groups at values of h negative to -15. Although FIGURE 3 leads to no definitive estimate of the number of imidazole groups, it is entirely consistent with the value 5 ± 1.

The evidence discussed above, which is based solely on an analysis of dissociation curves, leads to the following estimate of the numbers of the various ionizable groups in egg albumin:

a. Total cations................................ 41
b. Carboxyl minus guanidine.......................... 37
c. Amino...................... 22
d. Imidazole.......... 5
e. a-c-d = guanidine................................ 14
f. b+e = carboxyl.................................. 51

Egg albumin is reported to contain roughly 1 mol. of phosphoric acid. Such a group would be expected to contribute 2 eq. to the curve. One of these would be submerged in the carboxyl region close to pH 2 and the other in the imidazole region at about pH 7. Our estimates of carboxyl and imidazole groups will include any phosphoric acid present.

The ionizable groups in a protein may also be estimated from the amino acid content on the assumption that the protein is a simple polypeptide. Pertinent data from the literature on the amino acids of egg albumin have been used to arrive at the following estimate.

Amino Acid	Ionizable Group	Equi. per mol.
Arginine[9, 10]	Guanidine	14
Histidine[9]	Imidazole	4
Lysine[9]	Amino	15
Tyrosine[11]	Phenolic Hydroxyl	10
Glutamic Acid = 40[12]		
Aspartic Acid = 24[12]		
$\overline{64}$		
Amide Nitrogen 31[13]		
$\overline{33}$	Free Carboxyl Groups	33

[9] Vickery, H. B., & Shore, A. Biochem. Jour. 26: 1101. 1932.
[10] Vickery, H. B. Jour. Biol. Chem. 132: 325. 1940.
[11] Calvery, H. O. Jour. Biol. Chem. 94: 613. 1931–2.
[12] Chibnall, A. C. Personal communication.
[13] Shore, A., Wilson, H., & Stueck, G. Jour. Biol. Chem. 112: 407. 1936.

The most dependable of these determinations is, undoubtedly, that of arginine. The histidine analysis, moreover, can scarcely be so greatly in error as to raise a serious question of the presence of more than 4 mols. It is satisfactory, therefore, that the dissociation curves are entirely consistent with the established content of imidazole and guanidine compounds. The lysine content, on the other hand, is insufficient to account for the groups which react with formaldehyde. An alternative determination of free amino groups may be derived from the reaction of the protein with nitrous acid. The reaction of nitrous acid with egg albumin substantially confirms the formaldehyde titration since it gives a value of 23–24 amino groups.[5] There is some evidence[5] against the view that the discrepancy between the lysine content and the free amino groups is due to the presence of a amino groups and, so, there remains the possibility that diamino acids other than lysine are present.

There is a very great discrepancy between the free carboxyl groups indicated by the dissociation curves and those calculated from the amounts of glutamic and aspartic acids which have been isolated from egg albumin. This is probably due in part to incomplete recoveries of these two acids, although it may also be due to the presence of hydroxyglutamic acid or other dicarboxylic acids. No simple departure from polypeptide structure will suffice to explain the anomalies which have been discussed.

Finally, it will be observed that no allowance has been made for the contribution of the phenolic hydroxyl groups to the dissociation curve. This is due simply to the fact that the analysis of the curves has not extended above pH 9. It would be surprising if phenols dissociated significantly at so low a pH. There is, indeed, reason to doubt whether these groups contribute to the curve below pH 12. At the latter point, the value of h varies with μ but does not, even in high concentrations of KCl, materially exceed the value which it attains in formaldehyde at pH 9. That is to say, the protons dissociated from egg albumin between pH 9 and 12 in water do not exceed the number of amino groups present. If phenols have dissociated below pH 12, a corresponding number of amino groups must still retain their charge at this point. It seems improbable that amino cations should persist at such a high pH.

In the analysis of the dissociation curves which follows, we will assume that egg albumin behaves as a polyvalent ampholyte containing 51 carboxyl, 5 imidazole (possibly including 1 phosphate), 23 amino, and 14 guanidine groups.

EFFECT OF SALT

When the effect of [KCl] on the dissociation curve of egg albumin is examined, it is found to exhibit a very simple quantitative relation to h. Let us adopt the curve for 0.067 M KCl as the standard curve. At chosen values of h we derive from TABLE 2 the difference in pH, Δ pH$_s$, between the standard and one of the other curves and plot this difference as a function of h. A linear relation is disclosed over all buffered regions of the curves. It may be described by

$$\Delta \text{pH}_s = a_s h + \beta_s \qquad (1)$$

where a_s is the slope and β_s is the value of Δ pH at pH 4.9. The β term is merely the expression of the fact that the curves at different [KCl] do not intersect at the chosen Isoionic Point. For the curves of β-lactoglobulin over the same range of concentrations up to 0.67 M KCl a similar relation has been demonstrated. The values of a are very nearly the same for the two proteins, provided h be calculated for the same unit weight of protein. In the case of β-lactoglobulin, however, there is no β effect because all curves meet at the Isoionic Point. Since identical methods and assumptions were employed in the two studies, we do not believe that β_s can be dismissed as an artifact reflecting errors in the calculated L.J. potentials. At the same time, it should be observed that, even in the case of egg albumin, β is a quantity of significant magnitude only at the higher ranges of ionic strength.

The only observable effect of substituting KCl by MCl$_2$ at constant ionic strength seems to be a parallel displacement of the whole dissociation curve toward a lower pH to an extent increasing with μ. In other words, a_s is unaffected but β_s is changed in sign. The displacements due to Ca^{++} and Mg^{++} are almost identical and are slightly greater than those due to Sr^{++} or Ba^{++}.

The β correction is an empirical device for bringing all curves into coincidence at the Isoionic Point of the standard curve. Its justification is merely that the effect of salt on the curves can, then, be rationally related directly to the net charge of the protein. A more logical device would be to bring all curves into coincidence at pH 4.9,— the assumed Isoionic Point in the absence of salt. The correction term, then, becomes β_0 (TABLE 2) where $\beta_0 = \beta_s + 0.03$. The qualitative relation of β_0 to the nature of the ions present, and to their concentration, is consistent with a preferential combination of the protein with one of the ions of the salt which is present. In KCl the sign of

β_0 would then indicate a preferential combination with Cl^-, whereas in MCl_2 it would indicate preferential combination with the divalent cation. The specific effects of salts on electrophoretic mobilities and on membrane potentials in egg albumin solutions are capable of the same qualitative explanation. If specific ion association is involved, we should, however, expect the effect to be related in some way to the charge on the protein. Yet, we assume above that β_0 is independent of the charge. This is not an essential assumption. It would seem probable that β_0 represents the magnitude of the specific effects at the Isoionic Point and that these vary with h. In that case, a represents the resultant of the specific and non-specific effects insofar as they are related to the charge. The specific effects are, however, small compared with the general ion effects and so the observed values of a in different salts at the same ionic strength are not noticeably different.

On the assumption that a has the qualities of an activity correction, we have sought to relate it to $\sqrt{\bar{\mu}}$. A satisfactory linear relation up to 1 M KCl does emerge if we plot a_s as a function of the Debye-Huckel approximation for the activity coefficient of a univalent ion (Equation 6), using a plausible value for the "distance of closest approach." The simplest assumption with respect to the latter is to identify it with the apparent radius of the egg albumin molecule. For this we have adopted the value of 27.5 Å as recommended by Adair and others.[2] Values of Ln γ, designated Ln γ_1, calculated for $a = 27.5$ Å are given in TABLE 4. More precisely, however, we should expect a to exceed the egg albumin radius by an amount equal to the mean radius of the ions in the ion atmosphere. Now, in simple solutions of KCl, the observed activity coefficients of KCl may be reconciled with the Debye-Huckel equation if $a = 4$ Å[14] corresponding to a mean radius of the ions of 2 Å. We have, accordingly, calculated a second set of values for egg albumin, designated Ln γ_2, for the assumption that $a = 29.5$ Å. In FIGURE 4, a_s is plotted as a function both of Ln γ_1 and of Ln γ_2. The relations, up to $\mu = 1$, are linear and correspond with

$$- \alpha_s = 0.057 + 0.64 \text{ Ln } \gamma_1 = 0.060 + 0.7 \text{ Ln } \gamma_2.$$

[14] **MacInnes, D. A.** "The Principles of Electrochemistry" Rheinhold Publishing Corp. New York, 1939. p. 165.

TABLE 4

μ	α_s	$-\text{Ln } \gamma_I$	$-\text{Ln } \gamma_2$
0.0085	−0.0205	0.0584	0.0563
0.0167	−0.014	0.0688	0 0659
0.033	−0.007	0.0796	0.0757
0.067	−0.000	0.0895	0.0846
0.133	+0.005	0.0981	0.0922
0.267	0.009	0.1053	0.0986
0.667	0.014	0.1127	0.1050
1.19	0.015	0.1162	0.1080
2.38	0.016	0.1194	0.1108

μ	$\dfrac{\Delta \text{pH}}{\Delta h}$	$w_{\text{exp.}}$	w_I	w_2	$f = \dfrac{v_{\text{exp.}}}{w_2}$
0.0085	0.085	0.0585	0.0696	0.0717	0.82
0.0167	0.078	0.0505	0.0592	0.0621	0.81
0.033	0.071	0.0425	0.0484	0.0523	0.81
0.067	0.064	0.0345	0.0385	0.0434	0.80
0.133	0.059	0.0285	0.0299	0.0358	0.80
0.267	0.055	0.024	0.0227	0.0294	0.81
0.667	0.051	0.019	0.0153	0.0230	0.83
1.19	0.050	0.0185	0.0118	0.0200	0.93
2.38	0.049	0.0175	0.0086	0.0172	1.0

Curves of Sørensen, Linderstrøm-Lang, and Lund[4]

Ammonium Chloride					
0.05	0.0685	0.0395		0.047	0.84
0.20	0.0575	0.027		0.0319	0.84
0.99	0.049	0.017		0.0209	0.82
1.96	0.046	0.0135		0.0165	0.80
Ammonium Sulphate					
0.089	0.0610	0.031		0.0400	0.78
0.521	0.0510	0.0195		0.0246	0.81

Now, 0.057 and 0.060 are simply the values of 0.64 Ln γ_1 and of 0.7 Ln γ_2, respectively, when $\mu = 0.067$. Consequently,

$$\alpha_s = 0.64 \, \Delta \text{Ln } \gamma_1 = 0.7 \, \Delta \text{Ln } \gamma_2 \text{ and, from (1),}$$

$$\Delta \text{pH} - \beta_s = 0.64 \, h. \quad \Delta \text{Ln } \gamma_1 = 0.7 \, h. \quad \Delta \text{Ln } \gamma_2 \qquad (2)$$
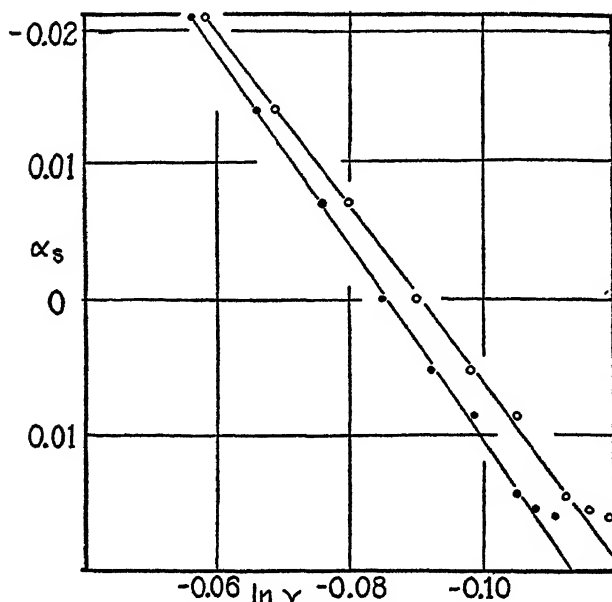
FIGURE 4.   Relation of α_s to $-\text{Ln } \gamma$.

O cf. Ln γ_1.
● cf. Ln γ_2.

Ionic Strength of Protein Ions

In protein solutions containing no simple ions other than those of the titrating agents, we may assume that the ionic strength (μ) is the sum of the contribution of the inorganic ions (μ_s) and that of the protein ions (μ_p). At chosen values of h in TABLE 3 we have calculated μ on the assumption that equation (2) applies. We may also calculate μ_s since $2\mu_s = \dfrac{g\,h}{45000} + [\text{H}^+] + q$, where $g = $ gm. protein per liter and $q = $ the eq. of acid or base necessary to bring g gm. of the protein preparation to the Isoionic Point. The difference, $\mu - \mu_s$, gives μ_p. The calculation is indirect and is based upon small differences in pH. Consequently, the results are highly irregular. They do, however, suggest that $2\mu_p$ does not, even at high values of h, substantially exceed the equivalent concentration of the net charge on the protein. We have suggested elsewhere[15] that μ_p does not vary much with the net charge and that the dipoles on the protein seem to make some contribution. TABLE 3 does not support this,

[15] Cannan, R. K.   Cold Spring Harbor Symp. Quant. Biol. 6.   1938.

though it does confirm the relatively small value of μ_p even at large values of h.

THEORETICAL

Let us restrict consideration to that section of the dissociation curve of a protein which lies between pH 2 and 6. The protein may, then, be treated as a polycarboxylic acid with a fixed cation charge of n and a variable net charge depending on the number of protons which have dissociated from the m COOH groups present. The successive proton dissociations may be defined by the constants k_1', k_2', . . . k_x' . . . k_m', such that

$$k_x' = \frac{[P_{n-x}] a_H}{[P_{n-x+1}]} = \frac{(P_{n-x})}{(P_{n-x+1})} \cdot \frac{\gamma_{n-x+1}}{\gamma_{n-x}} \cdot a_H \qquad (3)$$

where [ ] represents the concentration, ( ) the activity and γ the activity coefficient of the protein ion whose charge is denoted by the subscript, and a_H is the hydrogen ion activity. Applying the Mass Law to these m dissociations, we may derive the following expression for the relation of a_H to the number of protons dissociated, $(n - h)$:

$$n - h = \frac{n \cdot a_H{}^m + (n-1)k_1'a_H{}^{m-1} + \cdots + (n-m)k_1'k_2' \ldots k_m'}{a_H{}^m + k_1'a_H{}^{m-1} + \cdots + k_1'k_2' \ldots k_m'} \qquad (4)$$

We now assume that all of the COOH groups have the same intrinsic dissociating tendency defined by the intrinsic constant k_0. A relation between k_0 and k_x' may be sought in terms of electrostatic interactions. The first attempt to establish such a relation for a polyvalent ampholyte was made by Linderstrøm-Lang[16] on the basis of a very simple model. He considered the case of a spherical ampholyte in which the interactions of its charges with the dissociating protons and with the ion atmosphere could be represented by the electrostatic effect of a charge at the centre of the sphere equal to the net charge on the molecule. The validity of the working equation which Linderstrøm-Lang derived was restricted to the centre of the dissociation curve by reason of the assumptions made in its derivation. Moreover, the amphoteric characteristics which he assigned to his model were very different from those which are characteristic of proteins. In the treatment which follows, we have merely adapted the argument of Linderstrøm-Lang to the case of a protein having characteristics of the type we have attributed to the egg albumin molecule and extended

[16] **Linderstrøm-Lang, K.** Compt. rend. trav. lab. Carlsberg, **15**, No. 7. 1923–5.

its application to the whole course of the experimental curves. Following Linderstrøm-Lang, we may write

$$k_x' = \frac{m-x+1}{x} \cdot k_0 \cdot e^{2(n-x)b} \cdot e^{2(n-x+0.5)\,\mathrm{Ln}\gamma}$$

$$= \frac{m-x+1}{x} \cdot k_0' \cdot e^{2(n-x+0.5)w}$$

or $pk_x' = pk_0' - 0.87\,(n-x+0.5)w - \mathrm{Log}\,\dfrac{m-x+1}{x}.$ (5)

In these equations

$$b = \frac{\varepsilon^2}{2DrkT}, \qquad k_0' = k_0\,e^{-b}$$

$$-\,\mathrm{Ln}\,\gamma = \frac{\varepsilon^2}{2DkT} \cdot \frac{\kappa}{1+\kappa a} = b \cdot \frac{\kappa r}{1+\kappa a}$$

$$\text{and } w = b + \mathrm{Ln}\,\gamma = b\left[1 - \frac{\kappa r}{1+\kappa a}\right].$$ (6)

The term $\dfrac{m-x+1}{x}$ is the statistical expression of the multiple opportunities for the dissociation of the xth proton. The first exponential term in (5) represents the interaction of this proton with the net charge, $n-x$, of the protein. The final term is the Debye-Huckel expression for the activity coefficient ratio in (3).

The values of the constants k_1' . . . k_m' may be calculated from (5) for chosen values of k_0', w, m and n. These constants may then be introduced into (4), computations made of the values of h at chosen values of pH and a theoretical dissociation curve constructed for each value of w employed. A less tedious method is to derive values of w from the experimental curves and to compare these with those calculated from (6) using plausible values for a, r, and D. Linderstrøm-Lang has shown that the central sections of curves defined by (4) and (6) are approximately straight lines whose slopes are given by

$$\frac{\Delta pH}{\Delta h} = -\,0.87\left(w + \frac{2}{m}\right).$$ (7)

The slopes of the experimental curves between pH 3 and 4.5 have been determined graphically. They are recorded in TABLE 4 together with the values of w, $w_{\mathrm{exp.}}$, calculated from them $(m = 51)$. The slopes in the region of pH 11 have also been determined. They lead

to substantially the same values of $w_{exp.}$ ($m = 23$). In TABLE 4 and FIGURE 5, $w_{exp.}$ is compared with two sets of values of w calculated from (6). The series, designated w_1, is for the condition that $r = a = 27.5$ Å, and that designated w_2, for the condition that $r = 27.5$ Å and $a = 29.5$ Å. The value of D is assumed to 78.8. The linear relations which are apparent in FIGURE 5 correspond with

$$w_{exp.} = 0.007 + 0.74 \ w_1 = 0.8 \ w_2.$$

It is evident that $w_{exp.}$ is more simply related to w_2 than to w_1. In brief, if we assume that $r = 27.5$ Å and $a = 29.5$ Å, then the intro-
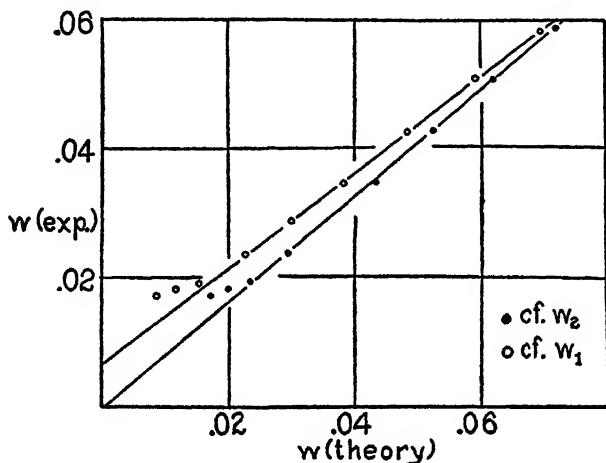
FIGURE 5.   Relation between experimental and calculated values of w.

duction of an empirical coefficient, f, into the interaction terms of (6) is all that is necessary to effect a quantitative reconciliation of the theory with the observed slopes of the curves at pH 3–4.5 and at pH 10–11. For ionic strengths up to 1, f has a value of 0.8. At higher ionic strengths, the value rises towards 1. In the calculations which follow, we will restrict consideration to curves for $\mu < 1$, substituting (9) for (5).

$$pk_x' = pk_0' - 0.7 \ (n - x + 0.5) \ w_2 - \text{Log} \cdot \frac{m - x + 1}{x}. \qquad (9)$$

It remains to be shown that (9) leads to consistent values for k_0' and that (9) and (4) suffice to describe the whole course of the experimental curves.

Values of k_0'

At the mid-point in the curve of a symmetrical polyvalent acid, $h = \dfrac{2n - m}{2}$ and $\mathrm{pH} = pk'_{\frac{m+1}{2}}$. It follows, from (9), that

$$\mathrm{pH_{mid.}} = pk_0' - 0.7 \left(\frac{2n - m}{2} \right) w_2.$$

Assuming the presence of 51 carboxyl, 5 imidazole, and 23 amino groups, the values of h at the mid-points of the contributions of these groups will be 15.5, -12.5, and -26.5, respectively. The corresponding values of $\mathrm{pH_{mid.}}$ are given in TABLE 5 together with the values

<div align="center">TABLE 5</div>

$$pk_0' = \mathrm{pH_{mid.}} - \beta_0 + 0.7 \frac{(2n - m)}{2} w_2$$

μ	w_2	β_0	COOH 15 5		Imidazole -12.5		Amino -26.5	
			$\mathrm{pH_{mid.}}$	pk_0'	$\mathrm{pH_{mid.}}$	pk_0'	$\mathrm{pH_{mid.}}$	pk_0'
0.0085	0.072	0.0	3.50	4.26	—	—	—	—
0.0167	0.062	0.01	3.63	4.27	—	—	—	—
0.033	0.052	0.02	3.74	4.29	7.23	6.75	11.07	10.07
0.067	0.0435	0.03	3.85	4.28	7.18	6.78	10.89	10.07
0.133	0.036	0.04	3.95	4.30	7.14	6.79	10.78	10.08
0.267	0.0295	0.07	4.03	4.28	7.08	6.75	10.69	10.07
0.667	0.023	0.10	4.14	4.30	7.06	6.75	10.58	10.01
1.19	0.020	0.14	4.21	4.32	—	—	—	—
2.38	0.017	0 22	4.30	4.33	—	—	—	—
Mean				4.29		6.7		10 07

Column 1 header: μ ; Column 2 header: w_2 ; with the $\dfrac{n-m}{2}$ ...

of pk_0' calculated from them. The latter are not inconsistent with the characteristics of the groups to which they are attributed, although that for the amino group is lower than would be expected.

Theoretical Dissociation Curves

The stoichiometric analysis of the experimental curves has indicated that the contributions of the three types of groups do not overlap to any important extent. By a simple approximation, therefore, (9) and (4) may be extended to cover the whole experimental range of pH. The equations are solved independently for each type of

group and the results are, then, combined to give the necessary data for the construction of a composite curve. Only in the region of pH 6 will such a curve depart significantly from the true curve for the ampholyte, owing to a slight overlapping of the carboxyl and imidazole constants. The computations are tedious but have been made in detail for $w_2 = 0.0435$, which corresponds with $\mu = 0.067$. The calculated values of h are compared, in TABLE 6, with a series

TABLE 6

THEORETICAL AND OBSERVED CURVES FOR $\mu = 0.067$

$w_2 = 0.043;\quad 0.7\,w_2 = 0.03;\quad \beta_0 = 0.03$

$pk_0'\,(COOH) = 4.29;\ pk_0'\,(imidazole) = 6.7;\ pk_0'\,(amino) = 10.07$

Calc. pH = Obs. pH − β_0	h	
	Calc.	Obs.
1.5	39.80	—
2.0	37.82	37.8
2.5	33 88	33.7
3.0	28.33	28.0
3.5	20.61	20.7
4 0	13.11	12.9
4.5	5.38	5.3
4.9	0 00	0.0
5 0	− 1.08	− 1.0
5.5	− 5.72	− 5.5
6.0	− 8 55	− 8.6
6.5	−10.67	−10.7
7.0	−12.28	−12.1
7.5	−13.57	−13.5
8.0	−14.4	(−14.0)
9.0	−15.68	−15.8
9.5	−16.76	−17.0
10.0	−19.22	−19.3
10.5	−22.91	−23.1
11.0	−27.78	−28.0

of observed values obtained by interpolation in the experimental curve. The remarkable agreement between them is good evidence that (4) and (9) describe the whole course of the curve with precision. A sufficient number of calculations have been made at other ionic strengths to satisfy us that the theory may be applied to the other curves with equal success.

The significance of the coefficient, f, is obscure. The most plausible explanation would seem to be that it represents the extent of departure of the egg albumin molecule from the assumptions of spherical symmetry and a random distribution of ionizing groups. It may, however, reflect errors in the assumptions respecting the values of the molecular weight and of a, r, or D. The coefficient could be raised to unity by reducing the anhydrous molecular weight to 36,000 or, less plausibly, by substantially increasing a, r, or D.

Two hypothetical curves are of some interest. One of these is the ideal curve for zero ionic strength. For this condition $w_2 = b = 0.128$. Data for the construction of this curve will be found in TABLE 7 which has been given the form shown in order to facilitate comparison with the experimental data of TABLE 2. The calculations were made in the same manner as those in TABLE 6. The curve was, then, drawn and the values given in TABLE 7 were obtained by interpolation. The second hypothetical curve is that for $w = 0$, or $\mathrm{Ln}\ \gamma_2 = -b = -0.128$. This condition could only be achieved at infinite ionic strength and, then, only if $a = r$. Alternatively, the curve may be referred to an ideal state in which the interaction terms are negligible,—a state which could be approached by a sufficient increase in a, r, or D. When $w = 0$, (5) reduces to

$$k_x' = \frac{m-x+1}{x} \cdot k_0'. \tag{11}$$

Now, von Muralt[17] and others have shown that, when this relation exists, the curve of the polyvalent acid becomes identical with that of m equivalents of a univalent acid having a dissociation constant equal to k_0'. Consequently,

$$\frac{a_H}{k_0'} = \frac{m-(n-h)}{n-h}. \tag{12}$$

The data for $w = 0$ in TABLE 7 were obtained by solving (12) for each type of group in turn and, then, combining the results into a composite curve. Now, it may be seen from (6), that the possible values of w must all lie between b and zero. The significance of TABLE 7 is that it defines the limiting shapes of the theoretical egg albumin curve. If the two limiting curves are compared, it will be found that their relative positions can be described by (2) with the same precision that this empirical equation describes the relations

---

[17] von Muralt, A. L. Jour. Amer. Chem. Soc. **52**: 3518. 1930.

<div align="center">

TABLE 7

THEORETICAL CURVES

</div>

μ	0	0 067	∞
w_2	0.128	0.043	0.00
$0.7\ w_2$	0.09	0.03	0.00
h	pH		
38	—	1.96	3.09
36	0.10	2.25	3.33
34	0.41	2.48	3.49
32	0.72	2.68	3.62
30	1.05	2.85	3.73
28	1.30	3.01	3.82
26	1.58	3.14	3.91
24	1.80	3.27	3.99
22	2.08	3.40	4.06
20	2.33	3.52	4.13
15	2.94	3.86	4.31
10	3.58	4.18	4.48
8	3.83	4.31	4.55
6	4.07	4.45	4.63
4	4.34	4.58	4.71
2	4.61	4.74	4.80
0	4.90	4.90	4.90
− 2	5.20	5.08	5.02
− 4	5.52	5.29	5.16
− 6	5.88	5.52	5.35
− 8	6.32	5.85	5.60
−10	6.90	6.32	6.03
−12	7.66	6.95	6.60
−14	(8.6)	(7.8)	(7.3)
−16	(10.1)	(9.2)	(8.7)
−18	10.85	9.77	9.25
−20	11.30	10.10	9.52
−22	11.68	10.38	9.71
−24	12.04	10 61	9.88
−26	—	10.82	10.03
−28	—	11.02	10.18
−30	—	11.25	10.34

between the experimental curves.   This equation is evidently a close approximation over the whole pH range to the requirements imposed by (4) and (9).   Now, according to (9), the change in pk_x' ($\Delta pk_x'$) with change in μ ($\Delta\ \mathrm{Ln}\ \gamma_2$) is:

$$pk_x' = 0.7 \ (n - x + 0.5) \ \Delta \operatorname{Ln} \gamma_2.$$

As a close approximation (*cf.*, equation (2)), therefore $\Delta pk_x' = \Delta$ pH when $n - h = x - 0.5$. This suggests an inquiry into the relation between $n - h$ and x when, at a given ionic strength, pH $= pk_x'$. The relation is a complex function of w, x, n, and m which reduces to simple terms only when $w = 0$ or when w is so large that the dissociation of successive protons are substantially separated on the pH scale. In the latter case $n - h = x - 0.5$ when pH $= pk_x'$. In the case of a molecule of the size of a protein, however, w can only have values which are relatively close to zero. When $w = 0$, we have, by combining (11) and (12) and setting $\alpha_H = k_x'$,

$$n - h = \frac{mx}{m+1} = x - \frac{x}{m+1}$$

That is to say, for no possible value of x can $n - h$ deviate from $x - 0.5$ by more than $\pm \dfrac{m-1}{2(m+1)}$, *i. e.* by more than ± 0.5. If we restrict consideration to the well-buffered region of the curve,—*i. e.*, to values of $n - h$ between 0.2 m and 0.8 m—then $n - h = x - 0.5$ ± 0.3. This is independent of the value of m. When w assumes small positive values, such as those attributed to egg albumin, the situation is not substantially altered, although the deviation of $n - h$ from $x - 0.5$ at a chosen value of h diminishes steadily as w increases. We conclude that almost the whole of a theoretical curve may be constructed with an error of less than 0.5 eq. simply by plotting calculated values of pk_x' as a function of $h = n - x + 0.5$. The laborious computations involved in the solution of (4) are thus avoided. Now, it is true that the probable errors in the titration of a single protein preparation should be less than 0.5 eq. It is doubtful, however, if the curve of a protein can be reproduced on independent preparations with a precision exceeding this. The graphic approximation which has been described can, of course, be improved for small values of w by using $h = n - x \left(\dfrac{m}{m+1} \right)$ in place of $h = n - x + 0.5$.

One of us[15] has recently shown that the carboxyl regions of the dissociation curves of egg albumin and of β-lactoglobulin may be described by the equation

$$\text{pH} = pk' - v \operatorname{Log} \frac{m - (n - h)}{n - h} \tag{13}$$

where k' and v both vary with the ionic strength. Kern[18] has independently used the same type of equation to describe the curves of certain polyacrylic acids. In the paper referred to, we showed* that $v - 1$ was related to $w \dfrac{m}{2}$ and suggested that (13) might be a more general approximation to the requirements of the Linderstrøm-Lang theory than is (7). This suggestion is withdrawn inasmuch as (13) tends to give somewhat low values for w when v is adjusted to give the best fit over a wide section of the observed curves. Equation (2) of the present paper is a much more comprehensive approximation than is either (7) or (13).

Effects of Ammonium Chloride and of Ammonium Sulphate

Sorensen, Linderstrøm-Lang, and Lund[4] have published careful studies of the effects of these salts on the curve of egg albumin between pH 4 and 5. From the slopes of their curves, we have derived $w_{exp.}$ with the aid of (13).** In TABLE 4 these values of $w_{exp.}$ are compared with those of w_2 calculated for the ionic strengths which they employed. The ratio $w_{exp.}/w_2 = f$ is seen to be very close to that which we have found for KCl and for the chlorides of the alkaline earths.

SUMMARY

1. Studies are reported on the effects of temperature, and of the concentration of protein and of various salts, on the dissociation curve of egg albumin.

2. The dissociation curve (pH 2-11) is consistent with the presence of 50-51 carboxyl, 4-5 imidazole, 23 amino, and 14 guanidine groups in a molecule of 45,000 grams anhydrous molecular weight. These values for imidazole and guanidine groups agree with the amounts of histidine and arginine which have been isolated from egg albumin. The amino and carboxyl groups both exceed by about 50% the number to be expected from the lysine and the dicarboxylic acid contents respectively.

18 Kern, W. Zeit. physik. Chem. A,181: 249. 1938.

* In the first paragraph of page 6 of reference 15 there are two misprints. h'/q should read $2h'/q$, and equation 9 should read pH $- p1 = (wq + 1) \, \text{Log} \dfrac{a}{1-a}$.

** The values of h reported by these authors were based on the assumption that egg albumin contains 14.28% nitrogen. The reason for this figure is not clear. Their values were recalculated for the generally accepted figure of 15.6% N and for a molecular weight of 45,000.

3. The slopes of the curves vary with the ionic strength and are substantially independent of the nature of the ions present. The apparent Isoionic Point depends on the nature of the ions as well as on their concentration.

4. The curves may be reconciled with Linderstrøm-Lang's application of electrostatic theory to proteins, provided two empirical corrections are applied. These are:

 a. the adjustment of the pH scale for each curve in such manner as to bring all curves to coincidence at the Isoionic Point.

 b. the introduction of a factor of 0.8 into the electrostatic interaction terms which, according to the theory, determine the slopes of the curves.

5. The question of the contribution of the protein ions to the apparent ionic strength is considered.

THE INFLUENCE OF pH ON THE MOBILITY AND DIFFUSION OF OVALBUMIN

By L. G. LONGSWORTH

*From the Laboratories of The Rockefeller Institute for Medical Research
New York City, New York*

INTRODUCTION

Some progress has been made in the development of a satisfactory theory for the electrophoresis of proteins including the correlation of mobilities with their other physical properties. Thus Abramson[1] has shown that in the region near the isoelectric point a proportionality exists between the mobility of the protein, as measured with the aid of the micro-cataphoretic method, and the amount of hydrogen ion bound to it as determined from the titration curve. To establish this proportionality he plots the pH as the abscissa for both the titration and mobility curves and shifts the former vertically until the acid bound by the protein is zero at the isoelectric point, *i. e.*, the pH at which the mobility is zero. The scale of ordinates for the mobilities is then selected to bring another point on this curve into coincidence with the titration curve and if the two curves then superpose, the proportionality is established. In most of the cases studied by Abramson and his collaborators the comparison has been made over a rather narrow pH interval in which the titration and mobility curves are approximately linear. It seemed desirable to extend the comparison over the entire pH-stability range of a typical protein the titration curve of which exhibits some characteristic feature that one might expect to be reproduced in the mobility curve. Ovalbumin is well adapted for this comparison since it has a wide stability range, pH 2.6–12.0 at 0° C., and has a characteristic plateau in its titration curve from pH 7.0 to 10.0. It is one of the purposes of this paper to report mobility measurements on ovalbumin over the pH interval from 1.8 to 12.8 and to compare these results with titration data on the same protein. The results have been interpreted with the aid of the Debye-Hückel-Henry theory. One of the uncertainties in the computations with this theory, *i. e.*, the friction coefficient of the protein, has been eliminated with the aid of diffusion measurements on ovalbumin over the same pH range and these results are also reported here. The

[1] Abramson, H. A. Jour. Gen. Physiol. **15**: 575. 1931–32.

behavior of ovalbumin outside of its stability range will also be considered briefly.

MOBILITY MEASUREMENTS

The mobility measurements to be reported in this paper have been made on the dissolved protein with the aid of the moving boundary method. The apparatus employed in these experiments is a modification of that used by Tiselius[2] and has been adequately described in previous papers from this laboratory.[3, 4, 5, 6] The computation of
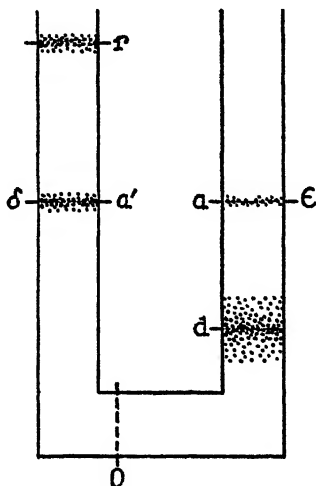

FIGURE 1.

mobilities from the electrophoretic patterns has not been discussed elsewhere in sufficient detail, however, and will now be considered.

It is a feature of the moving boundary method, as applied to proteins, that the rising boundary is usually sharper, and its position more easily determined, than the descending boundary, with the result that observations on the rising boundary have been used in preference to those on the descending boundary for the computation of mobilities. Dr. MacInnes and I showed,[4] however, that mobilities computed from observations on the rising boundary could be seriously

[2] Tiselius, A.   Trans. Faraday Soc. 33: 524. 1937.
[3] Longsworth, L. G., Cannan, R. K., & MacInnes, D. A.   Jour. Am. Chem. Soc. 62: 2580. 1940.
[4] Longsworth, L. G., & MacInnes, D. A.   Jour. Am. Chem. Soc. 62: 705. 1940.
[5] Longsworth, L. G.   Ann. N. Y. Acad. Sci. 39: 187. 1939.
[6] Longsworth, L. G., & MacInnes, D. A.   Chem. Rev. 24: 271. 1939.

in error if a rather involved correction for the δ effect were not made. Even in cases where mobilities have been computed from observations on the descending boundary the results may be in error if the complete concentration gradient curve through the boundary is not available and interpreted properly.   Thus the center of a schlieren band photograph of a boundary does not locate the position of the boundary correctly if the gradients in the latter are not symmetrical about the ordinate passing through the maximum value of the gradient. The location of a diffuse descending boundary is illustrated in FIGURES 1 to 3.   The electrophoresis cell, in which the boundaries between the buffer and protein solutions are formed initially at the level a–a',
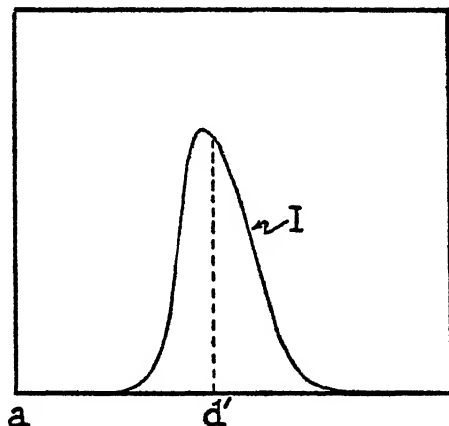
FIGURE 2.

is shown diagrammatically in FIGURE 1.   When formed the boundaries are quite sharp and their precise location presents no difficulty.   On the passage of a given quantity of electricity the boundary in the right hand side of the channel, for example, descends to a new position, d, but in so doing frequently becomes broad and diffuse as indicated by the shading in the figure.   The variation of the refractive index gradient, assumed proportional to the concentration gradient, through a typical boundary is shown in FIGURE 2.   The gradient-distance curve shown in this figure was actually obtained with a 0.5 per cent solution of ovalbumin in a 0.02 N sodium acetate–0.08 N sodium chloride buffer at pH 3.91.   Since the mobility is defined as the distance moved per second in a unit electric field by an average particle in the body of the protein solution the problem, then, is to locate d in such a manner that the boundary displacement, d–a, when multi-

plied by the cross-sectional area, A, of the channel and the protein concentration, $[P]$, will yield the same number of grams of protein as have simultaneously migrated through some reference plane, o— FIGURE 1, in the body of the protein solution. The latter quantity of protein, *i. e.*, $(d-a)A[P]$, is the quantity that is measured, except for a small volume correction to be considered later in this section, in a mobility determination by the Hittorf or gravimetric method. In order to locate d it is necessary to construct the concentration-distance curve shown in FIGURE 3, by integration of the gradient curve, and d is then the position of an ordinate, located by trial and error with the aid of a planimeter, for which the shaded areas on either side cancel.[7]
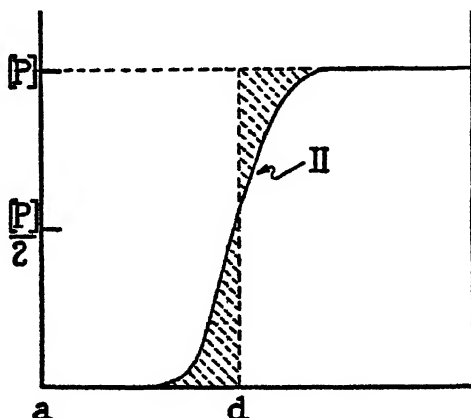
FIGURE 3.

The ordinate d will be recognized as the position the boundary would have if it had descended without loss of its original sharpness. It may be recalled[4] that the ϵ boundary, FIGURE 1, is due to gradients of buffer salt and not to protein, thereby differing from the δ boundary which includes both salt and protein gradients. Consequently the region above d has been swept free of protein by the current and, since the composition of the body of the protein solution remains unchanged on electrolysis, the quantity of protein, $(d-a)A[P]$, is equal to that which migrates through the reference plane o. Of course, this computation is valid only if the transport of protein is the result of electrolytic migration and is not complicated by transfer due to convective disturbances.

[7] *Cf.* **Tiselius, A.**   Nova Acta Soc. Sci. Upsala IV. **7** (4): 35.   1930

The moving boundary method has been criticized[8] because the potential gradients acting on the protein molecules in the boundary are not known. While this is true the demonstration just completed shows clearly that only the value of the gradient in the body of the protein solution is required for the mobility computation. Moreover, the mobility computed as outlined here corresponds to the pH and ionic strength of the protein solution and not to these properties of the buffer solution with which the former is in contact at the boundary.

The construction of the concentration curve of FIGURE 3 from the gradient curve of FIGURE 2 is rather laborious and it has been found,
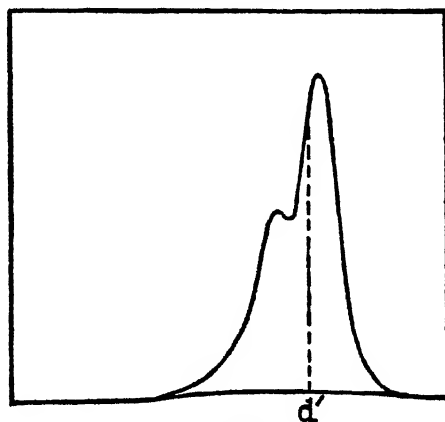
FIGURE 4.

for all of the cases to be considered in this paper, that the boundary position is given with sufficient accuracy by the ordinate, d' of FIGURE 2, that bisects the area under the gradient curve, *i. e.*, the level in the boundary at which the protein concentration is one half of its original value. Also, a complication arises in the computation of the mobility of crystalline ovalbumin in the pH interval from 4 to 10 owing to the presence in this material of two electrically separable constituents. This is illustrated by the refractive index gradient curve obtained at pH 5.33 and shown in FIGURE 4. Since comparison is to be made with titration data on ovalbumin of similar complexity, a weighted mean mobility of the two components was computed from the position of the bisecting ordinate shown at d' in FIGURE 4, for example.

⁸ Mukherjie, J. N.   Proc. Roy. Soc. London A103: 102.   1923.

VOLUME CORRECTION

In the moving boundary method the reference plane, *o* of FIGURE 1, is fixed with respect to the apparatus, whereas in a Hittorf experiment it is fixed with respect to the solvent. Consequently a correction, usually small, must be applied to the results obtained with the moving boundary method for volume changes in the system that cause solvent to move through the reference plane. These volume effects, discussed elsewhere in more detail,[4, 9] arise in part from the reactions occurring at the electrodes. If the correction is not made, one obtains a slightly different mobility, if for example, a cadmium anode is used instead of a silver anode.

Due to the lack of adequate density data for the buffer solutions used, no correction for the volume changes has been applied to the mobilities reported in this paper. Some computations from available data indicate that the error thereby introduced does not exceed 1 or 2 per cent in most instances. When the pH of the solution is within a few hundredths of the isoelectric pH, however, a considerable portion of the observed boundary displacement may be due to volume changes in the apparatus. Consequently the precise location of the isoelectric pH involves a careful evaluation of the volume correction. We are attempting to obtain experimental confirmation of the computed volume correction with the aid of moving boundary measurements on raffinose which is, presumably, electrically inert.[10] Although not completed this work has progressed sufficiently to say that the volume correction will not shift the isoelectric pH of ovalbumin at 0.1 μ and 0° from the value of 4.58, observed by both Tiselius and Svensson[11] and myself, by more than a few hundredths of a pH unit.

MOBILITIES OF OVALBUMIN

All of the mobilities have been measured in buffers of 0.1 ionic strength and with a protein concentration of about 0.5 per cent. The results of the mobility measurements over the pH-stability range of ovalbumin are recorded in the third column of TABLE 1. The composition of the buffers is given in the first column of TABLE 1 and the corresponding pH, at 0° C., in the second column. These solutions are not to be considered as buffer standards since no special precautions were taken in the preparation of many of the buffer salts for weighing. The values of all buffers below pH 9 were determined at room

⁹ MacInnes, D. A., & Longsworth, L. G.   Chem. Rev. **11**: 171.   1932.
¹⁰ Washburn, E. W.   Jour. Am. Chem. Soc. **31**: 322.   1909.
¹¹ Tiselius, A., & Svensson, H.   Trans. Faraday Soc. **36**: 16.   1940.

<div align="center">

TABLE 1

MOBILITIES OF OVALBUMIN AT 0° C. IN BUFFER SOLUTIONS OF CONSTANT
IONIC STRENGTH 0.1

</div>

1	2	3	4	5
Buffer	pH	$u \times 10^5$	$e \times 10^4$	e/u
0.1 N HCl–0.5 N glycine	3.05	6 25	5.30	8.48
0.02 N NaAc*–0.2 N HAc–0.08 N NaCl	3.62	3 89	3.39	8 71
0.02 N NaAc–0.1 N HAc–0.08 N NaCl	3.91	2.79	2.34	8 39
0.1 N NaAc–0.2 N HAc	4.34	1.10	0.82	7.45
0.1 N NaAc–0.15 N HAc	4.47	0.51	0.37	7.25
0.1 N NaAc–0.1 N HAc	4.64	−0.20	−0.15	7.50
0.1 N NaAc–0.02 N HAc	5.33	−2.82	−1.99	7.06
0.1 N NaAc–0.01 N HAc	5.65	−3.53	−2.52	7.14
0.02 N NaCac†–0.02 N HCac–0.08 N NaCl	6.12	−4 46	−3.11	6 97
0.02 N NaCac–0.004 N HCac–0.08 N NaCl	6.79	−5.16	−3.60	6 98
0.02 N NaV‡–0.02 N HV–0.08 N NaCl	7.83	−5.92	−4.04	6.82
0.1 N NaOH–0.2 N glycine	10.28	−6 21	−4 86	7 83
0.1 N NaOH–0.125 N glycine	10 88	−7.00	−5.69	8.13
0.1 N NaOH–0.1 N glycine	11.81	−9.31	−7.50	8 05
	Mean			7.63
	Average deviation, %			7.2

* Ac—acetate.
† Cac—cacodylate.
‡ V—diethyl barbiturate.

temperature with a glass electrode standardized against 0.05 M potassium acid phthalate as pH 4.00. The correction of the pH values of these buffers to 0° was assumed to be negligible. The pH values of the glycine-sodium hydroxide buffers were interpolated from the data of Sørensen[12] extrapolated to 0° C. In no case did the pH and ionic strength of the protein solution differ significantly from the corresponding properties of the buffer solution used as solvent.

It may be noted that all of the buffer solutions listed in TABLE 1 contain only monovalent ions. Some measurements have also been made in buffer solutions of the same ionic strength prepared from different salts and the results are given in TABLE 2. In this table the third column contains the directly measured mobility in the buffer solvent listed in the first column. In the fourth column are given mobilities interpolated from the data of TABLE 1 and a comparison of these with the directly observed values indicates that substitution,

[12] *International Critical Tables.* McGraw-Hill Book Co., Inc. 1: 83. 1926.

TABLE 2

SPECIFIC EFFECTS OF BUFFER IONS ON THE MOBILITY OF OVALBUMIN

1	2	3	4	5
Buffer	pH	$u \times 10^5$ obs'd	$u \times 10^5$ interpo- lated*	Differ- ence, %
0.1 N NaAc–0.5 N HAc	3.92	2.75	2.75	.0
0.02 N NaCac–0.1 N HCac–0.08 N NaCl	5.42	−3.07	−3.05	.6
0 0833 M NaH$_2$PO$_4$–0.0055 M Na$_2$HPO$_4$	5.59	−4.40	−3.42	28.6
0.0125 M NaH$_2$PO$_4$–0.0025 M Na$_2$HPO$_4$– 0.08 N NaCl	6.10	−4.71	−4.42	6.6
0 0625 M NaH$_2$PO$_4$–0.0125 M Na$_2$HPO$_4$	6.12	−4 94	−4.46	10.8
0.0025 M NaH$_2$PO$_4$–0.0025 M Na$_2$HPO$_4$– 0 09 N NaCl	6.71	−5.46	−5.10	7.1
0.025 M NaH$_2$PO$_4$–0.025 M Na$_2$HPO$_4$	6.80	−5.92	−5.17	14.5
0.1 N HCl–0.1 N glycine†	1.78	7.56⎫		1.8
0.02 N HCl–0.08 N NaCl†	1.77	7.70⎭		

* Interpolated from data of TABLE 1.
† Ovalbumin denatured in these buffers.

at constant pH and ionic strength, of one monovalent buffer salt for another, or of the neutral salt, sodium chloride, does not result in appreciable changes of mobility. On the other hand the change from a univalent buffer salt mixture to one containing the divalent HPO$_4$$^=$ ion has a decided effect on the mobilities. The results obtained with the last two buffers of TABLE 2 cannot be compared with the data of TABLE 1 since the pH values are below the stability range of ovalbumin. Here again, however, a considerable alteration in the composition of the monovalent buffers, at constant pH and ionic strength, produces only a slight change in the mobility of the resulting denatured ovalbumin.

TITRATION CURVE AND COMPARISON WITH THE MOBILITY DATA

Professor R. K. Cannan of the Chemistry Department of the New York University College of Medicine very kindly furnished me, in advance of its publication,[13] data from which the titration curve of ovalbumin at 0° C. and an ionic strength of 0.1 could be constructed. Most of the data were obtained by him at 25° C. but enough measure-

---

13 Cannan, R. K., Kibrick, A., & Palmer, A. H. Ann. N. Y. Acad. Sci. 41: 243. 1941

ments were made at 4° C. to demonstrate that the temperature co-
efficient of the acid portion of the curve is similar to that of the acetate
system while that of the basic portion agrees well with the temperature

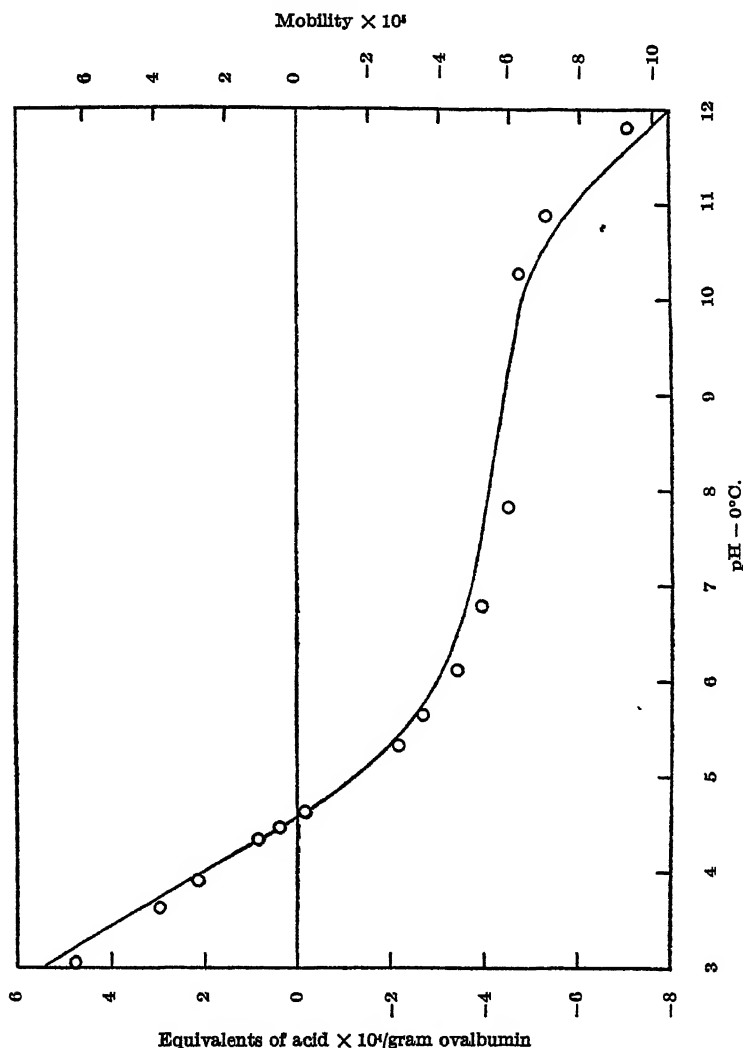FIGURE 5.

coefficient of the glycine-sodium hydroxide system. Consequently, in
plotting the titration curve of FIGURE 5, I have made the same tempera-
ture corrections for pH as in the case of the buffers of TABLE 1. In
this figure the pH values are plotted as abscissae and the number of

equivalents, e,[14] of acid bound per gram of protein as ordinates, this scale being shifted so that $e = 0$ at pH 4.58, the isoelectric point of ovalbumin at $\mu = 0.1$. This empirical adjustment of the titration curve is necessary if e is to measure the net charge on the protein. This appears to make allowance for ions other than hydrogen that may be combined with the protein at pH values near the isoelectric point. Using the ordinate scale given on the right hand side of FIGURE 5 the mobilities of TABLE 1 are indicated by the circles, and it will be seen that these parallel the titration curve rather closely. The constancy of the proportionality between the mobility and the number of equivalents of acid bound by the protein may be determined as follows. Values of e were interpolated from the titration curve for the pH values listed in TABLE 1 and are recorded in the fourth column of that table. From these data and the corresponding mobilities, values of the proportionality constant, e/u, were computed and are recorded in the last column of TABLE 1. The average deviation of this ratio from the mean value of 7.63 is 7.2 per cent and, although this exceeds the apparent experimental error of either the mobility or titration values, the ratio is sufficiently constant to warrant the conclusion that, except for minor secondary effects, the mobility of ovalbumin is proportional to the number of equivalents of acid bound by the protein at constant ionic strength at any pH within its stability range.

Since a proportionality between the mobility and the acid bound by the protein would be expected only if the friction coefficient of the latter were independent of the pH, it appeared desirable to obtain additional evidence on this point. This coefficient may be derived from diffusion measurements and these have been made, as will be described in the following section, at the same temperature and in the same buffer solutions as the mobility determinations.

DIFFUSION MEASUREMENTS

The schlieren scanning method used in our electrophoresis work is also adapted for recording the refractive index gradients that arise in the diffusion of proteins. Moreover, acting upon a suggestion of Dr. D. A. MacInnes of these Laboratories, the standard Tiselius electrophoresis cell has proved to be satisfactory for diffusion measurements. In practice, boundaries are formed initially in the two sides

---

[14] It may be noted that e, as defined here, is related to the quantity h used by Cannan, Kibrick, and Palmer[15] by the relation $h = Me$. In this equation M is the molecular weight of the protein, taken as 45,000 in the case of ovalbumin.

of the channel and are shifted into view from behind the opaque horizontal plates of the cell as in an electrophoresis experiment. The center sections of the cell are then isolated by displacement to one side and diffusion patterns, *i. e.*, schlieren scanning photographs of the refractive index gradients, are obtained at intervals until the diffusion process has approached the closed ends of the sections. For the 1.4 per cent protein solutions used in this work the time required is four or five days at the thermostat temperature of 0° C. In order to avoid monopolizing our equipment for this period with a single determination it has proved practicable to have several diffusion experiments, together with an electrophoresis measurement, proceeding simultaneously. The diffusion cells are supported in the thermostat in such a manner that they can be successively moved into the focus of the camera without disturbing the diffusion process. Zero time for the diffusion is taken as the time at which the boundaries are formed. Values of the diffusion coefficient computed on this basis do not exhibit a drift with time, from which it seems reasonable to conclude that the process whereby the boundaries are shifted into view does not disturb the diffusion. Moreover, diffusion patterns obtained immediately before and after the movement of the cell into and out of the camera focus indicate no disturbances from this source.

The use of the electrophoresis cell as a diffusion cell yields diffusion patterns in duplicate since boundaries are present in both sides of the channel. One method for utilizing these patterns for the computation of the diffusion coefficient is shown in FIGURE 6. In the case of ideal diffusion the maximum value of the refractive index gradient, $\left(\dfrac{dn}{dx}\right)_m$, varies inversely as the square root of the time, t, according to the relation[15] $\left(\dfrac{dn}{dx}\right)_m = \Delta n / \sqrt{4\pi D t},$ in which D is the diffusion coefficient and Δn is the refractive index increment. Hence a series of ordinates are drawn on a sheet of coordinate paper at values of the abscissae equal to the reciprocals of the square roots of the times at which patterns were obtained. An enlarged image of the pattern is then traced on the paper after the position of the latter has been shifted to bring the base line of the pattern into coincidence with the axis of abscissae and the proper ordinate passes through the maximum value of the gradient. If diffusion has been normal these maxima fall, as illustrated in FIGURE 6, on a straight line passing through the origin. The diffusion coefficient is then computed from the slope, m,

[15] Lamm, O. Nova Acta Soc. Sci. Upsala IV. 10 (6). 1937.

of this line and the areas under the curves, the latter being proportional to the refractive index increment Δn, according to the formula $D = \frac{1}{4\pi}\left(\frac{\Delta n}{m}\right)^2$. The areas, A, under each of the curves are indicated
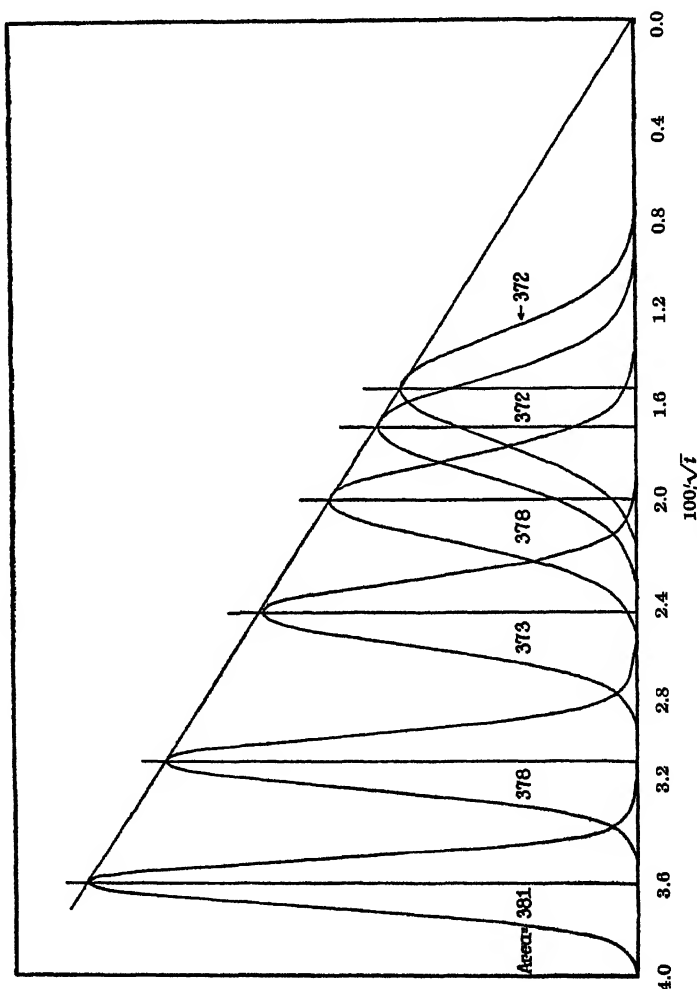

FIGURE 6.

in FIGURE 6 and it will be noted that they are essentially constant as required by the simple theory.[15]

The diffusion coefficients of ovalbumin at 0° C., determined as outlined above, are recorded in TABLE 3. The composition of the buffers and their pH values are given in the first and second columns, respec-

tively. The relative viscosities of these buffers were also measured at 0° C. and are recorded in the third column of the table. In columns 4 and 5 are given values of the diffusion coefficients, corrected for the viscosity of the buffer solution used as solvent, as obtained from the two sides of the channel respectively. The last column contains the averaged values of the corrected diffusion coefficient. It may be noted that at each pH within the stability range the average fluctuation of the individual values is about 2 per cent whereas the variation of the averaged values with the pH is somewhat greater, about 4 per cent, and shows a regular increase with increasing pH. If additional dif-

TABLE 3

DIFFUSION COEFFICIENTS OF OVALBUMIN AT 0° C.

1	2	3	4	5	6
Solvent	pH	η	$D\eta \times 10^7$	$D\eta \times 10^7$	$D\eta$ (average) $\times 10^7$
0.1 N HCl–0.1 N glycine	1.77*	—	(2.70–	→1.90)†	—
0.1 N HCl–0.2 N glycine	2.43*	1.012	3.82	3.75	3.78$_5$
0.1 N HCl–0.5 N glycine	3.05	1.041	3.88	4.04	3.96
0.1 N NaAc–0.1 N HAc	4.64	1.045	3.96	4.02	3.99
0.02 N NaV–0.02 N HV–0.08 N NaCl	7.83	1.019	4.14	4.09	4.11$_5$
0.1 N NaOH–0.1 N glycine	11.81	1.028	4.10	4.17	4.13$_5$
0.1 N NaOH–0.075 N glycine	12.81*	—	(2.78–	→2.12)†	—

* Denaturation occurred during diffusion.
† Observed values in both channels decreased during period of observation as indicated.

fusion measurements on ovalbumin, or a recomputation of the present results using properties of the refractive index gradient curves other than the area and maximum ordinate should confirm the small increase in the diffusion coefficient with pH observed in this research it would indicate a slight consolidation of the molecule as it loses some fifty-eight protons in the transition from pH 3.0 to pH 12.0. However, since the variation with pH is but slightly greater than the probable error of the measurements, this variation will be neglected in the computation, to be considered later in this paper, of a value for the ratio, e/u, with the aid of the Debye-Hückel-Henry theory.

Although the diffusion measurements of Polson[16] were made at 20° C. it is of interest to note that his value for a 1.4 per cent solution

[16] Polson, A. Kolloid Zeit. 87: 149. 1939.

of ovalbumin at pH 4.64 is 4.00×10^{-7} when corrected to 0° C. with the aid of the relation $D_0 = D_{20} \dfrac{273}{293} \dfrac{\eta_{20}}{\eta_0}$. This is in good agreement with the value, 3.99×10^{-7}, reported in TABLE 3 for the same pH and protein concentration.

COMPUTATION OF e/u

The theory required for the computation of a value for e/u may be reviewed briefly as follows. An isolated particle carrying a charge, q, in a dielectric medium very quickly attains a constant velocity, v, when an external field of strength, F, is applied. This velocity is proportional to the charge and to the field strength, *i. e.*,

$$v = \frac{1}{f} qF$$

in which the factor of proportionality, $1/f$, is the reciprocal of the friction coefficient, f, of the particle. The velocity in unit field, *i. e.*, v/F, is the mobility, u, so that

$$u = \frac{q}{f}. \tag{1}$$

In an actual mobility determination, however, the particle is not isolated but is surrounded by an ion atmosphere. The effects of this atmosphere are (1) to screen the particle from the full effect of the external electric field and (2) to modify the viscous flow of solvent past the moving particle. Both effects depend upon the ionic strength of the solution. The screening effect was computed by Debye and Hückel[17] to be $\dfrac{1}{1 + \kappa a}$ for a spherical particle, in which a is the distance of closest approach of the ions of the atmosphere to the particle and κ is the reciprocal of the "thickness" of the ion atmosphere. The flow modification, also for a spherical particle, was computed by Henry[18] to be a function, $\varphi\,(\kappa a)$, which becomes 1 for small values of κa and $3/2$ for large values of this product. For spherical particles dissolved in a solution of finite ionic strength equation 1 thus becomes

$$u = \frac{q}{f} \frac{\varphi(\kappa a)}{1 + \kappa a}. \tag{2}$$

The friction coefficient, f, is given by the relation

[17] Debye, P., & Hückel, E. Physik. Zeit. **24**: 305   1923.
[18] Henry, D. C. Proc. Roy. Soc. London **A133**: 106.   1931.

$$f = \frac{kT}{D}$$

in which k is the molecular gas constant and T the absolute temperature. Substitution of this value of f in equation 1 gives

$$u = \frac{Dq}{kT}$$

a relation that is valid at zero ionic strength quite independently of the shape of the particle. At finite ionic strength, however, equation 2 must be employed and the factors in that expression correcting for the effect of the ion atmosphere have been evaluated in sufficient detail only for spherical particles. Hence it is necessary at this stage in the development to introduce the approximation that the particles are spherical and obey Stokes' law, *i. e.*,

$$f = 6\pi\eta r \tag{3}$$

in which η is the viscosity of the medium and r the distance from the center of the particle at which solvent slippage begins. Elimination of f between equations 2 and 3 then gives

$$u = \frac{q}{6\pi\eta r} \frac{\varphi(\kappa a)}{1 + \kappa a}. \tag{4}$$

If the acid bound by the particle at a given pH is combined strictly as hydrogen ion

$$q = eM\epsilon$$

in which M is the molecular weight of the substance and ϵ is the electronic charge. Consequently

$$u = \frac{eM\epsilon}{6\pi\eta r} \frac{\varphi(\kappa a)}{1 + \kappa a} \tag{5}$$

and

$$e/u = \frac{6\pi\eta r}{M\epsilon} \frac{1 + \kappa a}{\varphi(\kappa a)}. \tag{6}$$

In computing with equation 6 for the case in question, *i. e.*, $\mu = 0.1$ and 0° C., M for ovalbumin has been taken as 45,000[11] and a has been taken as equal to $r = kT/6\pi\eta D = 27.8$ Å. On this basis $e/u = 4.60$, a value differing by 40 per cent from the experimentally determined figure of 7.63. Alternatively, one can compute a molecular weight from the experimental value of e/u with the aid of equation 6 and

obtain $M = 27,200$, a value that cannot be reconciled with the results of other methods of determining this constant. The conclusion must be drawn, therefore, that although a proportionality exists between the mobility and the amount of acid bound by the protein at a given pH the proportionality constant does not agree with that given by the theory outlined here.

DISCUSSION

Several explanations for the discrepancy between the observed and computed values of e/u have been suggested. Gorin[19] has made some computations of (1) the effect of the finite size of the ions of the atmosphere upon the screening factor and (2) the effect of deviations from a spherical shape upon both the screening factor and the flow modification. Gorin's correction for the finite size of the electrolyte ions does not appear to be very significant at the ionic strength used in this research. The shape of the particle, on the other hand, is important,—an ellipsoid undoubtedly having a lower mobility than a sphere of the same volume and charge. Although the use of a radius computed from the diffusion coefficient corrects for deviations from a spherical shape at zero ionic strength, the term, $\left(\dfrac{\varphi\,(\kappa a)}{1 + \kappa a}\right)$, in equation 5 reduces the limiting mobility by a factor of 0.285 at the ionic strength of 0.1 used in this research and this term is valid only for spherical particles.

Other possible explanations for the discrepancy are (1) that a portion of the acid bound by the protein is in the undissociated form and (2) that the protein forms weak, *i. e.*, incompletely dissociated, salts with the buffer ions. As a matter of fact Tiselius and Svensson[11] criticize the use of titration data for computing the charge on the protein because of these possibilities. They obtained excellent agreement with the theory outlined above in their study of the variation of the mobility of ovalbumin with ionic strength at the constant pH of 7.10 using Adair and Adair's[20] membrane potential measurements to compute the charge. It is of interest in this connection to compare the charge on ovalbumin at pH 7.10, as a function of the ionic strength, as obtained from the membrane potentials of Adair and Adair in sodium phosphate buffers with that obtained from the titration data of Cannan, Kibrick, and Palmer[18] in the presence of potassium chloride. The values are recorded in TABLE 4. In computing the

[19] Abramson, H. A., Gorin, M. H., & Moyer, L. S. Chem. Rev. **24**: 345. 1939.
[20] Adair, G. S., & Adair, M. E. Trans. Faraday Soc. **36**: 23. 1940.

1	2	3	4
μ	Valence from membrane potentials	Valence from titration curve	Difference
0 01	−12 0	−13 8	1 8
0 02	−12 6	−14 7	2 1
0 05	−13 1	−16 0	2 9
0 2	−15 1	−19 1	4 0

valence from the titration data, column 3, the shift in the isoelectric point with ionic strength, as measured by Tiselius and Svensson, was taken into account. It will be noted that the valences estimated from the titration data are greater than those computed from the membrane potentials and that the differences, column 4, increase with increasing ionic strength. These differences are in agreement, if the influence of the protein upon the activity coefficients of the buffer ions is neglected, with Dr. Steinhardt's point of view[21] that the protein forms weak salts with potassium and sodium ions, in addition to hydrogen ions, on the alkaline side of the isoelectric point.

The author does not feel, however, that the results of Tiselius and Svensson establish the validity of either the theory reviewed here or the use of membrane potentials for the computation of the charge. These investigators used phosphate buffers and the specific effects of these buffers have already been noted in connection with TABLE 2. Their experimental confirmation of a theory developed for spherical particles would seem to imply that the ovalbumin molecule is spherical, and yet other evidence[22] exists that it is not. Finally, the work of Hitchcock[23] and Failey[24] indicates that the protein does alter the activity coefficients of the buffer ions and this effect should be considered in computing the charge on the protein from either membrane potentials or titration data.

[21] Steinhardt, J. Ann. N. Y. Acad. Sci. **41**: 287. 1941.
[22] Neurath, H. Jour. Am. Chem. Soc. **61**: 1841. 1939.
[23] Hitchcock, D. I. Jour. Gen. Physiol. **16**: 357. 1932–33.
[24] Failey, C. F. Jour. Am. Chem. Soc. **54**: 2367. 1932.

OBSERVATIONS ON OVALBUMIN OUTSIDE OF ITS
STABILITY RANGE

Although only a few observations on ovalbumin outside of its pH stability range have been made by the author, these were sufficient to suggest that the electrophoretic method could be used to advantage in a study of the denaturation process. At the suggestion of Dr. A. E. Mirsky of the Hospital of The Rockefeller Institute, insolubility in dilute buffers at pH 4.6 has been taken as the criterion for denatured ovalbumin. Portions of each of the ovalbumin solutions used in mobility or diffusion measurements were tested for denaturation and it is significant that the appearance of even a slight precipitate on bringing the sample to pH 4.6 invariably coincided with some abnormality in the diffusion and electrophoretic patterns. The effect of progressive denaturation on the apparent diffusion coefficient has already been noted in TABLE 3. Some of the abnormalities in the electrophoretic patterns are illustrated in FIGURES 7 and 8 and these will now be considered.

The patterns in FIGURE 7a were obtained in a 0.1 N NaOH–0.075 N glycine buffer at pH 12.81 after dialysis of the ovalbumin for 72 hours. The patterns of FIGURE 7b were obtained under similar conditions except that the protein was dialyzed for only 8 hours prior to electrophoresis. In both cases the symmetry between the patterns for the rising and descending boundaries is reasonably good and both patterns yield a mobility of -10.3×10^{-5} for the fast component and $-7.4_5 \times 10^{-5}$ for the slow component. The relative amount of the fast component is much greater, however, in the sample that had been exposed to the alkaline buffer for 8 hours than in the sample exposed for 72 hours. This suggests that the fast component may be the native protein and the slow component some form of denatured ovalbumin. The large mobility difference between the two forms found in this instance is in contrast with the observation that completely heat, or acid, denatured ovalbumin has mobilities at pH 6.80 and 10.28 but slightly less than the native form and appears to be quite homogeneous electrically.

The patterns of FIGURE 8 were obtained in a 0.1 N HCl–0.2 N glycine buffer at pH 2.44 and the rising and descending boundaries are very far from being mirror images of each other. Asymmetries of this type have not yet been explained. The correct interpretation of these electrophoretic patterns, obtained with protein systems in which denaturation is occurring, is of considerable importance and will doubtless contribute to our knowledge of the factors determining the stability of proteins.
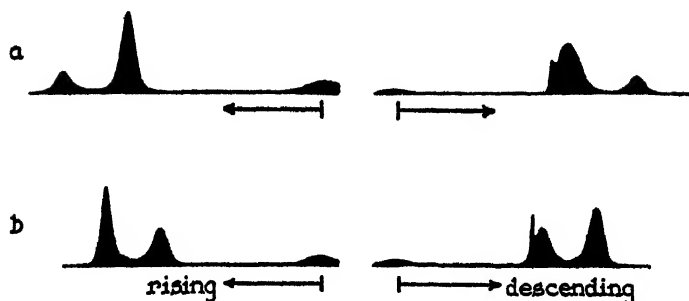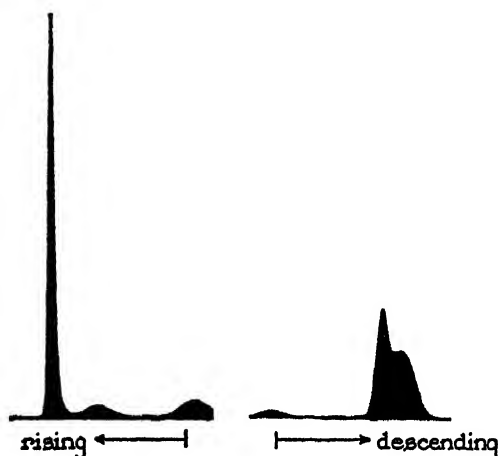
FIGURE 7.

FIGURE 8.

ACKNOWLEDGEMENT

PARTICIPATION OF ANIONS IN THE COMBINATION OF PROTEINS WITH ACIDS

By Jacinto Steinhardt

From the Research Laboratory of the Textile Foundation at the National Bureau of Standards

By making use of a number of special advantages offered by the insoluble protein, wool keratin, it has proved possible to study in detail a number of hitherto obscure phenomena which are closely related to the amphoteric properties of proteins. Although, as shall be shown presently, most of the phenomena to be described are probably not restricted to wool, this insoluble material of doubtful chemical individuality is especially well suited for an investigation of the equilibria involved, essentially free of extraneous complications. Thus, (a) because wool is insoluble the amounts of acid and base with which it combines can be determined by direct titration of the solutions in which it is immersed, and no assumptions as to relations between activity and concentration need be made in calculating the amounts bound; and (b) because wool has almost exactly equivalent numbers of amino and carboxyl groups and only a very small content of histidine, the titration curve with acid is unusually simple, being determined by the properties of the carboxyl groups (and amino groups) as a whole and not by an indeterminate fraction of them. The analysis of the results obtained is thereby greatly facilitated, (c) unlike most native proteins, the state of wool is not affected by exposure to acid in the pH range which is usually investigated. When proper precautions are taken the acid titration curve represents a strictly reversible phenomenon, and may be justifiably analyzed in terms of equilibrium equations, (d) the effects of salt on the titration curve of wool are especially large and obey a particularly simple law, as has been shown elsewhere.[1]

Largely as the result of simplifications which were made possible by the first three properties just described, progress has been made toward understanding the fourth, and this in turn has led to the analysis of specific anion effects described in the present paper.

---

[1] Steinhardt, J., & Harris, M. Jour. Research Nat. Bur. Standards **24**: 335. 1940. RP1286.

Evidence has been presented in two previous papers in support of
the view that the combination of wool with hydrochloric acid may
be resolved into two partial reactions, one with hydrogen ions and
one with chloride ions, and that each reaction is governed by its own
equilibrium constant.[1, 2] On the basis of this assumption and ap-
plication of the law of mass-action to the association of both hydrogen
ions and chloride ions with the protein, it proved possible to account
quantitatively for the large effects produced by various concentrations
of potassium chloride on the curve relating combination of the protein
with hydrochloric acid to pH. Among other consequences of this
view was a strong possibility that the specific affinities for wool of the
anions of different acids might vary considerably and that therefore
the position of the titration curve of this and other proteins with
respect to the pH axis might vary by correspondingly large amounts,
according to the acid used. The existence of such a variation in the
positions of titration curves would constitute critical evidence in sup-
port of the hypothesis of stoichiometric anion association, since alter-
native analyses of the salt effect, previously considered, such as those
based upon the Donnan membrane equilibrium equations,[1] offer no
obvious basis for the prediction of such an effect, nor for its corollary,
combination to unequal extents with each of two or more anions
present in mixtures.

The present paper describes measurements of the combination of
wool with nineteen different acids, ranging in familiarity and com-
plexity from some of the mineral acids most commonly used in the
laboratory through the simpler aromatic sulfonic, carboxylic, and
phenolic acids to a soluble mono-azo acid dye. It is shown that very
wide differences do indeed exist between the positions on a pH grid
of the titration curves of the same protein obtained with different
strong acids and that these differences may be ascribed to wide varia-
tions in the anion dissociation constants characterizing the correspond-
ing protein-anion combinations. The quantitative aspects of these
differences are shown to be in full agreement with generalizations of
the equations already presented to take into account the effects ob-
tained with chloride ions alone. Predictions as to the effects of
variations of anion concentration and of temperature, based upon
these equations, have been tested and confirmed in detail. As a result
it has proved possible to establish a quantitative measure of the
relative affinities of the anions of various strong acids for proteins

[1] Steinhardt, J., Fugitt, C. H., & Harris, M.   Jour. Research Nat. Bur. Standards 25:
519   1940.   RP1348.

on the basis of a small number of measurements of acid combined and pH.

In order to demonstrate that the existence of differences in affinity between the ions of different acids is in no way restricted to wool, or even to the insoluble class of proteins for which the theory of anion-association was specifically proposed,[1] measurements of the combination of a number of the same acids with the most familiar and well-characterized of the *soluble* proteins, crystalline egg albumin, have also been made. Since comparable differences in the positions, with respect to the pH coordinate, of the titration curves obtained with different acids are found with both proteins, it is clear that the hypothesis of anion association is relevant to both. Thus current interpretations of the titration curves of proteins, soluble as well as insoluble,[3-9] and of all the colligative phenomena which depend on the combination of protein with acid or base, may require enlargement.

The acids employed in the present research were selected from among those commercially available, or those most easily prepared, with the purpose of obtaining data which could be simply calculated and interpreted in terms of molecular size, chemical structure, and the effects of specific substituents. For this reason most of the results reported in the present paper have been obtained with acids which are monobasic and virtually totally dissociated in water. A limited number of weak acids which are analogues of some of the strong acids used (*e. g.*, benzoic acid is analogous with benzenesulfonic acid in the sense used here) have been included for the specific purpose of examining the effects of this analogy.[10] Likewise it has not been possible to exclude dibasic acids entirely in experiments designed to determine the effect of certain substituents which undergo partial acidic dissociation in the pH range investigated, but an effort has been made to exclude dibasic acids, such as sulfuric acid and many of the

[1] Cohn, E. J., Green, A. A., & Blanchard, H. H. Jour. Am Chem. Soc **59**: 509. 1937.

[4] Hitchcock, D. I. Cold Spring Harbor Symp. Quant. Biol **6**: 24. 1939.

[5] Hitchcock, D. I. In Schmidt, C. L. A., "Chemistry of the Amino Acids and Proteins" p 596. Charles C Thomas, Springfield, Illinois. 1938.

[6] Kekwick, R. A., & Cannan, R. K. Biochem. Jour. **30**: 227. 1936.

[7] Loeb, J. "Proteins and the Theory of Colloidal Behavior." McGraw Hill, New York 1922.

[8] Schmidt, C. L. A. In "Chemistry of the Amino Acids and Proteins." p. 720. Charles C. Thomas, Springfield, Ill. 1938.

[9] Speakman, J. B., & Stott, E. Trans. Faraday Soc. **31**: 539. 1934; Trans. Faraday Soc **31**: 1425. 1935.

[10] Extensive results obtained with a comparable number of weak acids are not reported here. Work still in progress shows that no fundamental modification of the views expressed in this paper are required by the data obtained with weak acids, although a complete analysis of the latter results involves a number of additional factors. **Steinhardt, J., & Harris, M.** Proc. Am. Soc. Biological Chemists, Jour. Biol. Chem. **134**: xcvii. 1940.

disulfonic acids, in which the second dissociation constant is so large as to make difficult a distinction between combination of the wool with singly and doubly charged anions.   An effort was also made to include organic acids with all of the more common dissociating groups. Since strong acids containing phenolic or carboxylic radicals as dissociating groups can only be obtained by loading with substituents such as halogens or nitro groups, other related acids have been included in order to make possible an independent appraisal of the effect of various substituent groups, and of more strictly physical factors such as the size and shape of the anion produced.

EXPERIMENTAL PROCEDURE

The methods employed in this series of investigations have been described elsewhere.[1, 2, 11]

RESULTS AND DISCUSSION

Results Obtained in the Absence of Salts

The results obtained with sixteen different strong acids in the absence of salt at 0° C. are assembled in FIGURE 1.   In drawing the curves an effort has been made to represent the measurements fairly with continuous curves rather than to preserve the homogeneity of appearance of the data.   For this reason certain neighboring curves show a tendency to cross or to converge in certain limited regions. In most cases the tendency is not larger than can be accounted for by experimental error, or by the uncertainties inherent, especially with acids of high affinity, in the application of a small correction for alkaline ash.[12]   In measurements with a number of acids, the reversibility of the combination has been tested by subjecting the fibers to prolonged washing, by replacement of the combined anions on immersing the fibers in solutions of other anions having higher affinities, and by approaching the equilibrium from both sides.

The most noticeable feature of the data assembled in FIGURE 1 is the wide range of positions with respect to the pH coordinate which the titration curves occupy.   It is also apparent that, except for the minor deviations noted above, the curves form a coherent family, near neighbors following fairly parallel courses.   The change in the

[11] Steinhardt, J., Fugitt, C. H., & Harris, M.   Jour. Research Nat. Bur. Standards 26: 293.   1941.   R.P.1377.
[12] Sookne, A., Fugitt, C. H., & Steinhardt, J.   Jour. Research Nat. Bur. Standards 25: 61.   1940.   RP1314.
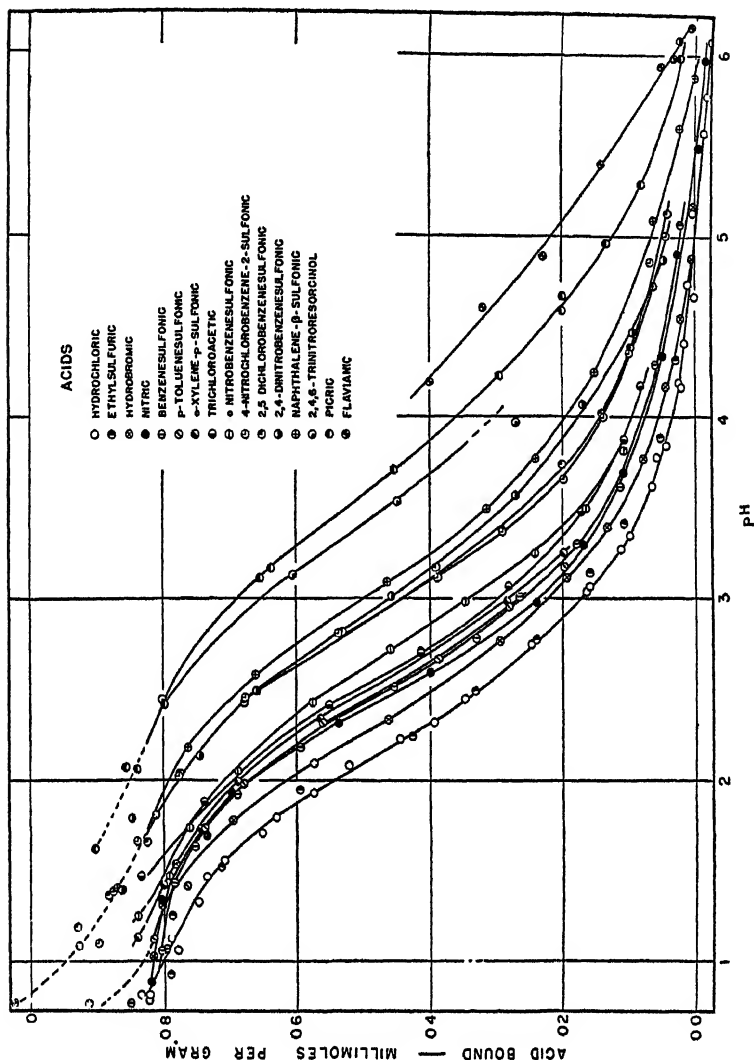
FIGURE 1. Combination of wool protein with sixteen different strong acids as a function of pH at 0° C.

The measurements with flavianic acid are represented in milli-equivalents instead of millimoles. Points representing the combination of less than 0.14 millimole per g of three acids (p-toluenesulfonic, o-xylene-p-sulfonic, and o-nitrobenzenesulfonic) would fall at lower pH values than are required for congruence with the curves for the other acids, and have been omitted to avoid confusing the figure. Data for the two dibasic acids, flavianic and trinitroresorcinol, are only partially represented, as explained elsewhere.

Tables of the data included in FIGURES 1, 4, 5, 6, 7, 8 are published elsewhere. (Jour Research Nat. Bur. Standards 26: 293. 1941.)

slopes of the lower portions of the curves, from the more sharply inflected curve for hydrochloric and ethylsulfuric acids at one end of the pH range to the least sharply inflected curves for picric and flavianic acids at the other end, tends to be gradual and progressive. The appearance of the curves as a group is entirely consistent with the view that they all describe analogous phenomena.

It is a familiar fact that the position on a pH grid of the titration curve of the common weak bases is determined by their own dissociation constants, and not by the dissociation constant or by other chemical properties of the acid employed in the titration. Sets of titration curves of the same base obtained in dilute solutions with any number of different acids, weak or strong, are expected to be practically superimposed on one another. Exceptions arise, as in the work of Cannan and Kibrick[13] and of Greenwald[14, 15] with carboxylic acids and phosphates, when the compound formed by the reaction of the base with certain acids is not typically salt-like and does not exist in solution predominantly in the dissociated form. This exception finds its counterpart, in the case of ampholytes, in the formation of only partially dissociated stoichiometric complexes between the anions of the acids and the ampholyte which has combined with hydrogen ions. When these are formed, the position of the titration curves, and to a lesser extent their shapes, will differ according to the extent to which the anion is dissociated.[1] An equation which, within certain special restrictions, shows the dependence on K_A', the anion dissociation constant, of the amounts of acid combined at any concentration of acid (or of hydrogen ions and of the conjugate anions) has already been given.[1]

The wide range of the measurements represented in FIGURE 1 serves to emphasize the very large differences which may be found when different acids are used. If the position of each curve is characterized by the pH value at which half the maximum amount of acid (about 0.4 millimole per gram) is taken up, there is a difference of almost 2 pH units between the curves shown at the extremes of the series; even wider differences between titration curves obtained with other acids at higher temperatures have been found. Between these extremes, the results obtained with the other acids represented are distributed with a fair degree of uniformity, with no evident tendency for many to resemble closely the results obtained with hydrochloric

[13] Cannan, R. K., & Kibrick, A. Jour. Am. Chem. Soc. **60**: 2314. 1938.
[14] Greenwald, I. Jour. Biol. Chem. **124**: 437. 1938.
[15] Greenwald, I., Redish, J., & Kibrick, A. C. Jour. Biol. Chem. **135**: 65. 1940.

acid, hitherto most widely used in measuring the acid-combining properties of proteins.[16]

Attention must also be directed to another prominent feature of many of the sets of data represented in the figure. This is the combination of amounts of acid in excess of 0.82 millimole per g, the amount reported as the maximum in the work with hydrochloric acid. It was earlier pointed out that this "maximum" of 0.82 millimole per g was sometimes exceeded slightly when high concentrations of hydrochloric acid ($> 0.2\ M$) were used. Measurements of acid combined are susceptible to a relatively large experimental error when such high concentrations of acid are present, because (a) only a small proportion of the acid initially present is combined, (b) some acid hydrolysis of the protein cannot be avoided, and (c) exchanges of water between the partially hydrated fibers and the solutions, difficult to evaluate exactly, have a relatively large influence on the results obtained. For these reasons relatively little weight was placed on these measurements, although they were consistent with the results of similar determinations on bone and hide collagen,[17] and might be accounted for by the presence in proteins of extremely feeble basic groups, such as the amide groups of glutamine and asparagine, and the large content of peptide nitrogen. Special interest, therefore, attaches to the finding that in every case in which sufficiently high concentrations of other acids are used, a second step in the curve of acid combination apparently begins at pH values more acid than those at which the curve that covers the usually investigated range of combination flattens. Since this fairly well-marked "maximum" of about 0.82 millimole per g has been shown to correspond closely to the primary amino content of the fibers, the additional uptake of acid in more concentrated solutions must therefore be accounted for by reaction with the more weakly basic groups, undissociated in the fiber because they are too weak to act as proton acceptors toward carboxyl groups, or else to combination by means of entirely different mechanisms.

The displacement of the pH range in which the excess combination with most of the acid is found removes much of the uncertainty attending the earlier observation of similar excess combination with hydrochloric acid. All three of the possible sources of experimental error mentioned above are reduced or eliminated, since the propor-

[16] Comparable differences between the titration curves of wool with hydrochloric and with picric acids were found by Elöd, E., Trans. Faraday Soc. **29**: 377. 1933.

[17] **Beek, J., Jr.** Jour. Research Nat. Bur. Standards **14**: 217. 1935. RP765; **21**: 117. 1938. RP1119.

tion of the total amount of acid initially present which is combined is then much larger. In addition, the excess found in concentrated solutions (1.03 — 0.82 = 0.21 in the case of naphthalenesulfonic acid) is far outside the limits of experimental error or of the range of effects on concentration due to exchange of water between fibers and solution.

Although the scatter of the points representing the combination of more than 0.8 millimole per g is considerable, the portions of the curves which represent the second step of acid combination appear to run parallel to one another to about the same extent as do the main curves at higher pH values. The displacements of the curves thus appear to correspond, qualitatively at least, to the similar displacements of the parts of the curves which represent the combination of the first 0.82 millimole per g. Should further work establish that the concepts applied in this paper to the lower sections of the curves may also be employed in the analysis of the excess combination of acid, the possibility is opened of studying in an aqueous environment the combination of these weakly basic groups with acid, and thus determining the number of these groups and the extent of their dissociation.

Calculation of the Affinity of Anions for Wool

Any simple analysis of the experimental results, in terms of the hypothesis of stoichiometric anion association, depends on the assumption that the curve of acid combination is homogeneous, *i. e.*, determined by the combination of hydrogen ions and of anions by sets of essentially similar groups in the fiber. If this is not the case, it is without meaning to attempt to ascribe a shift in the position of the titration curve as a whole to differences in single dissociation constants. The assumption is empirically justified when the displacement in the positions of the curves affects all parts of the curves to the same extent.

With minor exceptions noted elsewhere the data assembled in FIGURE 1 appear to justify a treatment based upon such homogeneity. In order to respect the condition of homogeneity, essential to a simple analysis of the data, the treatment that follows is applied only to the part of the titration data which represent amounts of acid bound lower than 0.8 millimole per g.
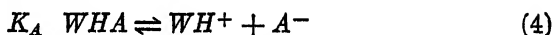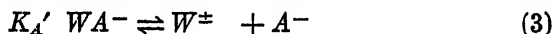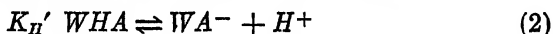
RELATION OF THE POSITIONS OF THE TITRATION CURVES TO THE ANION AFFINITY

Since it has been assumed that the differences in the positions of the curves with respect to the pH axis are determined by differences

in K_A', the protein-anion dissociation constant, it is appropriate to attempt to evaluate K_A' for the anions of each of the acids used by means of the position of the corresponding titration curve.[18] It has already been shown that changes in the positions of the titration curves obtained with a single acid, in the presence of pre-determined concentrations of the anions of the acid added in the form of neutral salts, can be described by the equation:

$$\text{Fraction of wool combined} \equiv \frac{[WHA] + [WH^+]}{[WHA] + [WH^+] + [WA^-] + [W^{\pm}]}$$

$$= \frac{1}{1 + \dfrac{K_H'}{a_H}\left[\dfrac{a_A + K_A'}{a_A + K_A}\right]} \qquad (1)$$

in which a_H and a_A represent the activities of the hydrogen-ion and of the anion respectively, and the constants K_H', K_A', K_H, and K_A govern the following hypothetical dissociation equilibria:

$$K_H' \quad WHA \rightleftharpoons WA^- + H^+ \qquad (2)$$

$$K_A' \quad WA^- \rightleftharpoons W^{\pm} \; + A^- \qquad (3)$$

$$K_A \quad WHA \rightleftharpoons WH^+ + A^- \qquad (4)$$

The constant K_H governing a fourth possible equilibrium:

$$WH^+ \rightleftharpoons W^{\pm} + H^+ \qquad (5)$$

does not appear in equation (1), since it is obvious that the four constants are interrelated by the relation $K_H' K_A' = K_H K_A$.

The assumptions entering into the definition set down in the left-hand member of equation (1) and the relations between the magnitudes of the several dissociation constants have already been discussed (36). It has also been shown that each experimental curve is actually somewhat broader (covering a wider range of pH values) than the S-shaped curves described by equation (1), but that the displacement between curves obtained at different values of a_A are accurately described by the equation.[19] The difference between the curves obtained

---

[18] The reciprocal of K_A' is referred to hereafter as the anion affinity constant, or simply as the affinity.

[19] Reasons have been given for attributing this broadening of the experimental curves to the high degree of polyvalency of the wool protein.[1] The difference between theoretical curves for monovalent acids and bases and experimental curves of analogous polyvalent substances can generally be taken into account by introducing an empirical fractional exponent on the hydrogen ion term in the mass-law expression for the monovalent acid. Kern, W. Zeit physik. chem. 189A: 249. 1938; Biochem. Zeit. 301: 338. 1939. The present simplified treatment also necessarily leaves out of account all effects due to the mutual interaction of dissociating groups in the molecule except for the effect implied broadly in the inequalities $K_H' < K_H$ and $K_A < K_H'$.

with and without a constant concentration of anions present is also accurately described by converting equation (1) into the equivalent form, valid when salt is absent:

$$\text{Fraction combined } = \frac{[WHA] + [WH^+]}{[WHA] + [WH^+] + [WA^-] + [W^\pm]}$$

$$= \frac{1}{1 + \dfrac{K_H{}'}{a_H}\left[\dfrac{a_H + K_A{}'}{a_H + K_A}\right]}. \quad (6)$$

The derivation of equations (1) and (6) made clear that their validity is limited to conditions under which the term $[WH^+]$ is greater than the term $[WA^-]$. These conditions are realized, except when much salt is present, when $K_A{}'$ is greater than K_H. The latter constant has not been evaluated, but it must be considerably larger than $K_H{}'$, which has been assigned the value 6.3×10^{-5} at $0°$ C.[1] Thus equation (6) cannot be used for anions having affinities which result in a displacement of the titration curve of more than 1 or 2 pH units to the right of the position of the curve obtained with hydrochloric acid.

It is possible to derive an equation, similar to equation (6), which has the complementary condition for validity, *i. e.*, $[WA^-] > [WH^+]$, or $K_A{}' < K_H$. This equation, applicable to acids of high affinity for wool, is obtained by equating the fraction combined to

$$\frac{[WHA] + [WA^-]}{[WHA] + [WA^-] + [WH^+] + [W^\pm]}$$

instead of to

$$\frac{[WHA] + [WH^+]}{[WHA] + [WH^+] + [WA^-] + [W^\pm]}$$

and leads to the general result:[20]

$$\text{Fraction combined } = \frac{1}{1 + \dfrac{K_A}{a_A}\left(\dfrac{a_H + K_H}{a_H + K_H{}'}\right)} \quad (7)$$

or, in the absence of salt, to the equivalent form:

---

[20] When salt is present, this equation gives the amount of acid *plus salt* combined since the excess negative charge due to $[WA^-]$ may be neutralized in part by adsorbed cations other than hydrogen ion. The amount of salt combined is large only when anions of high affinity are present or when high concentrations are used, so that the titration curve falls in a region of high pH values.

$$\text{Fraction combined} = \cfrac{1}{1 + \cfrac{K_A}{a_H}\left(\cfrac{a_H + K_{II}}{a_H + K_H'}\right)} . \tag{8}$$

The dependence of the pH coordinate of the mid-point of each curve on K_A' or K_A, in terms of either equation (6) or (8), may be described by equating the right hand member of each equation to 0.5. When this is done with equation (6) the result is

$$K_A' = \frac{a_H (a_H + K_{II})}{K_H'} - a_H = \frac{a_H^2 - a_H K_H'}{K_H'} \cdot \frac{K_{II}}{K_{II} - a_H} \tag{6'}$$

With equation (8) the result is

$$K_A' = \frac{K_{II}}{K_H'} \cdot a_H \left(\frac{a_H + K_H'}{a_H + K_{II}}\right) \tag{8'}$$

Equations (6') and (8') both contain constants which have not previously been evaluated (K_A and K_H) as well as the constant K_H' to which a value has previously been assigned.[21] Due to electrostatic requirements, however, $K_{II} > K_H'$ and $K_A' > K_A$.[1] The ratio between pairs of dissociation constants which differ mainly because the electrostatic effect of a single charge (such as the first and second dissociation constants of symmetrical dicarboxylic acids) always lie between 10 and 1000. If this analogy is accepted K_H probably lies between $10^{-3.2}$ and $10^{-1.2}$. Since the factor $\dfrac{K_H}{K_H - a_H}$ in equation (6') must be positive, the curves obtained with acids of lower affinity than hydrochloric acid (not reported in this paper) further restricts K_H to values greater than $10^{-2.1}$. The factor thus approaches unity for values of a_H at the mid-points of practically all the curves represented in FIGURE 1. Thus, between limits represented by the mid-point of the hydrochloric acid curve and the mid-points of curves for which the condition of validity for equation (6') ($K_A' > K_H$) no longer applies, the equation may be simplified to

$$K_A' \doteq \frac{a_H^2 - K_H' a_H}{K_H'} . \tag{6''}$$

The corresponding simplification of equation (8'), valid when $a_H < K_H$ (throughout practically the entire range of pH values in which the mid-points of the curves occur) is

---

[21] In all the calculations that follow the reasonable assumption is made that K_H' is an intrinsic property of the protein, and is not appreciably affected by the nature of the acid combined with it.[18]

$$K_A' \doteq \frac{a_H{}^2 + K_H'a_H}{K_H'}. \tag{8''}$$

The calculation of K_A', the reciprocal of the affinity, thus depends in either case on an accurate knowledge of K_H', the intrinsic acidity constant of the carboxyl groups of the fiber. By assigning the previously determined approximate value at $0°$ C., 6.3×10^{-5} to this constant, the relation of K_A' to the pH of the mid-point has been calculated for this temperature according to both equations. The result of this calculation is plotted logarithmically in FIGURE 2, in which the mid-point pH, and therefore the value of the affinity constant, according to both equation (6'') and equation (8'') has been indicated for each of the acids represented in FIGURE 1.

It is apparent that in the region of low pH values both equations give practically identical results. The curves approach the same straight line having a slope of 2. In this region the ratio of affinities of two anions is given with a high degree of approximation by the antilogarithm of twice the difference in pH between the mid-points of the titration curves obtained with the acids corresponding to the anions.

At pH values above 3, however, the curves diverge appreciably, and at pH values above 4.2 (pK_H' at $0°$ C.), no real values of K_A' are given by equation (6''). Obviously, however, the conditions under which equations (6) and (6') are valid ($K_A' > K_H$) are no longer realized when the curves are displaced to such high ranges of pH. Since both equations lead to essentially the same relation between K_A' and mid-point pH in the only range of pH in which equation (6'') is nearly exact, all calculations of affinity in this paper have been based on equation (8''), regardless of the range of pH in which the midpoint of a given titration curve may lie.

It should be noticed that as the affinity increases, the difference in the position of the curve produced by a given increase in affinity becomes larger until Δ pH, which at one extreme is equal to $- 1/2$ $\Delta pK_A'$ at the other extreme approaches $- \Delta pK_A'$. As the latter limit is approached closely, the affinity constant becomes independent of the value assigned to K_H' and is equal to the mid-point a_H.

Included in FIGURE 2 is a third curve, also representing equation (8''), in which the value 9.32×10^{-5} has been used for K_H' instead of the value 6.3×10^{-5}. This curve shows the relation between mid-point pH and log K_A', at a temperature of $25°$,[2] and also illustrates the relatively large effect of variations in K_H' on the estimation of
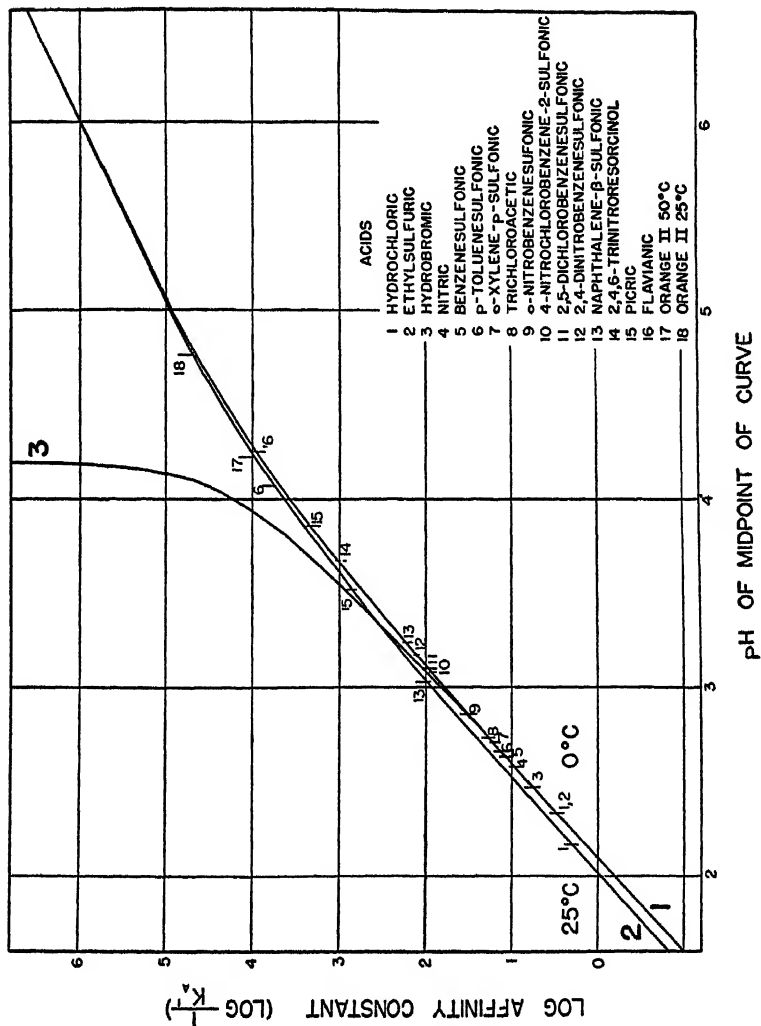
FIGURE 2. Relation between the position of the acid titration curve with respect to pH and the affinity of the anion of the acid for protein.

Curves 1 and 2 represent equation (8″) with numerical values of K_H' at 0° C. and 25° C. respectively inserted. Curve 3 represents equation (6″) with the numerical value of K_H' at 0° C. Curves 1 and 3 coincide at low pH values.

K_A' when the latter is large, and the virtual absence of this effect when the latter is small.

By making use of the curve which represents equation (8″) at 0° C., values of K_A' which correspond to each of the acids represented in FIGURE 1 have been determined. These values are given in the 5th column of TABLE 1. The ion with least affinity, chloride, has a value of K_A', 0.35; picrate, the most tightly bound monovalent ion studied at this temperature, has an affinity almost one thousand times as great as that of chloride; and the divalent flavianate ion has an affinity still higher.[22] It is apparent from the results summarized in the table that the affinity of the anion tends to rise as the molecular weight increases, but the tendency is not regular and exceptions in the series occur.

Since the shapes of the experimental curves are broader than those predicted by the much simplified equations (6) and (8), the use of these equations in analyzing the shifts in position of the curves, when the anion (or its concentration[1]) is changed, would be rendered more significant if the progressive change in the slopes of the titration curves could also be shown to be a consequence of the same differences in K_A' which result in the displacements to progressively higher ranges of pH values. This demonstration has been attempted by substituting representative numerical values in equation (8). For the sake of definiteness, the value, 0.01, has been assigned to the ratio K_H'/K_H which has already been shown to have a range of reasonable values between 0.01 and 0.001. The result of this calculation with several different values of K_A' is shown in FIGURE 3. The calculated curves simulate the curves drawn through the experimental data not only in the relation between the magnitudes of their displacements from one another, but also in the relation of their slopes. The qualitative agreement with experiment of these calculations would not be altered by the assignment of other reasonable numerical values to the ratio K_H'/K_H.

ALTERNATIVE METHOD OF CALCULATING RELATIVE ANION AFFINITIES

In the 6th column of TABLE 1 each value of K_A' has been expressed as the ratio of the affinity of the corresponding anion to that of chloride ion. This has been done because the same relative affinity factor may be calculated directly from the experimental results represented in

---

[22] For purposes of comparison, the affinity of the divalent flavianate ion is expressed in the same units as the affinities of the other monovalent acids in the table.

TABLE 1

CALCULATION OF ANION AFFINITY CONSTANTS $(1/K_A')$ AT 0° C.

Acid	Molecular Weight	pH of Midpoint	$-\log K_A'$ from eq. 8"	K_A'	Affinity relative to aff. of chloride ion	pH at which curve crosses HCl 0.1 Molal curve	Affinity relative to chloride ion (method of crossing point)
Hydrochloric	36.5 ⎫	2.33	0.45	0.35			
Ethylsulfuric	126.1 ⎭						
Hydrobromic	80.9	2.47	0.74	0.18	1.9		
Nitric	63.0	2.58	0.96	0.11	3.2		
Benzenesulfonic	158.2	2.63	1.05	0.089	3.9	1.71	3.9
p-Toluenesulfonic	172.2	2.66	1.10	0.080	4.4	1.71	3.9
o-Xylene-p-sulfonic	186.2	2.71	1.21	0.062	5.6	1.73	4.1
Trichloroacetic	163.4	2.73	1.25	0.056	6.2	1.83	5.2
o-Nitrobenzenesulfonic	203.2	2.86	1.50	0.032	11.0	1.93	6.6
4-Nitrochlorobenzene-2-sulfonic	237.6	3.08	1.93	0.0118	29.7	2.05	8.6
2, 5-Dichlorobenzenesulfonic	227.1	3.09	1.95	0.0112	31.2	2.68	36.8
2, 4-Dinitrobenzenesulfonic	248.2	3.17	2.09	0.0081	43	2.68	36.8
Naphthalene-β-sulfonic	208.2	3.24	2.23	0.0059	59	2.82	51
2, 4, 6-Trinitroresorcinol	245.1	3.67	3.02	0.00096	366	3.05	86
Picric	229.1	3.86	3.34	0.00046	758		
Flavianic	314.2	4.24	3 94*	0 000115*	3020*		

* Calculated, for purposes of comparison, as if the anion were monovalent. Since equilibrium was not attained in 23 days, the relative affinity reported is minimal.
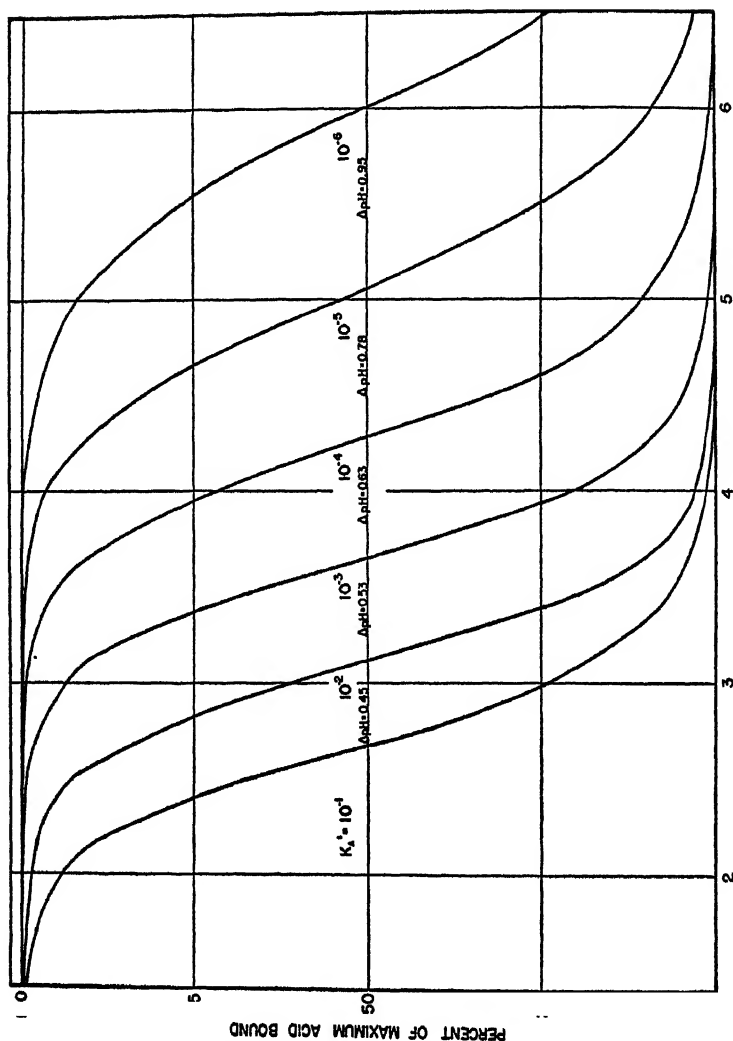
FIGURE 3.   Dependence of the slope of theoretical titration curves on the protein anion dissociation constant.   The two left-most curves are calculated from equation (6) and the others from equation (8).   The change from equation (6) to equation (8) between $K_A' = .01$ and $K_A' = .001$ has been made in order to preserve consistency with the assumption $\dfrac{K_H'}{K_H} = .01$.

FIGURE 1 by combining them with the data for hydrochloric acid at constant chloride concentrations in an earlier paper.[1] The values so obtained do not depend on the equations just used.

This alternative method of calculation is made possible by the fact that curves obtained in the presence of constant anion concentrations are less steep than those obtained in the absence of salt and therefore cover a wider range of pH values. Thus, a curve representing results obtained with hydrochloric acid at a constant concentration (0.1 M) of chloride ions would intersect all of the curves represented in FIGURE 1 with the exception of the curve for flavianic acid. The intersections with the curves for hydrobromic acid and picric acid would occur so near the extremes of the 0.1 M curve that they are not well suited to the present calculation, but all of the other intersections may be used. At the point of intersection of the 0.1 M hydrochloric acid curve with any curve in FIGURE 1 (two of the crossing points utilized may be seen in FIGURE 4), the same amounts of two different acids are bound by the wool at the same pH. The same degree of combination with two different anions is thus brought about, in the presence of different amounts of the respective anions, by virtue of the difference of affinity for wool characterizing these anions. It follows from the range of ratios of K_H'/K_H which has been adopted elsewhere that the ratio of the activities of the two anions in the two different experiments is nearly inversely proportional to the ratio of the respective anion affinity constants. The anion affinity ratios which are thus obtained should be the same as those given in column 6 of TABLE 1.

The activity of potassium chloride at 0.1 M at 0° is 0.0768.[23] The assumption has been made that the activity of the anion in each experiment made in the absence of salt is equal to the activity of the hydrogen ion. It is very unlikely that this assumption is strictly true, for the experimental value, pH, is related to both activities, and methods of obtaining individual ion activities are at present unknown. By dividing 0.0768 by the antilogarithm of the negative of the pH, the values given in the last column of TABLE 1 have been obtained. In view of the fact that the intersecting curves were drawn free hand and that the two kinds of data were obtained with different lots of wool the agreement of the values in the last column with those in column 6 is satisfactorily close.[24]

[23] Harned, H. S., & Cook, M. A. Jour. Am. Chem. Soc. 59: 1290. 1937.
[24] A third method of calculating relative affinities has also been made use of. Mixtures of acids have been equilibrated with wool and the relative amounts of the two acids com-
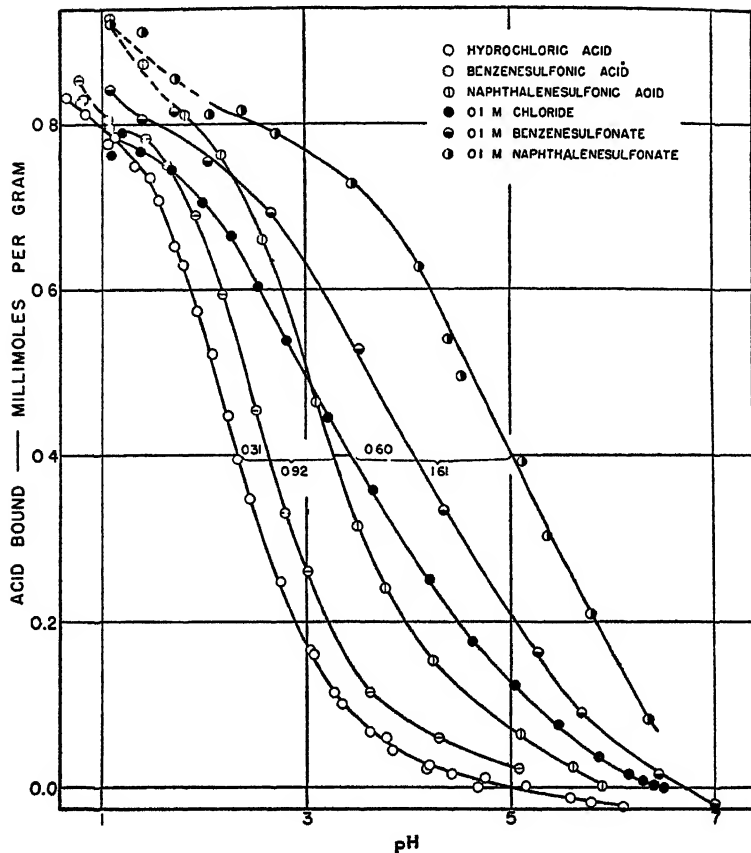
FIGURE 4.   Comparison of the titration curve of wool protein obtained with three different acids in the absence of salt at 0° C., with the curves obtained at the same temperature in the presence of a constant concentration (0.1 *M*) of the anions of the acids.

Results Obtained in the Presence of Salt

It should be possible to predict the displacements that would be found if the same acids were used in the presence of predetermined constant concentrations of their anions, added as neutral salts.   By starting with equations (1) and (7) instead of equations (6) and (8), and making the same simplification which was carried out in arriving at equations (6″) and (8″) formulas which give the relation between

bined have been determined.   The relative affinities of the two anions calculated from this ratio agree closely with the results of the methods described in this paper.   Since the method has thus far been applied to experiments with acids for which no data are given in the present paper, and since work with these acids is still in progress, no examples of these calculations are given here.

K_A' and the mid-point pH when the measurements are made in the presence of any constant concentration of anions is obtained:

$$K_A' = a_A \frac{a_H - K_H'}{K_H'}. \tag{1''}$$

$$K_A' \doteq a_A \frac{a_H + K_H'}{K_H'}. \tag{7''}$$

In this equation, the quadratic term in a_H, found in equations (6'') and (8''), does not appear. In the range of relatively low affinities, therefore, a given change in K_A' will, in the presence of a constant anion concentration produce twice the pH change which corresponds to this change in K_A' in the absence of salt. With anions having higher affinities, however, displacements in mid-point pH values obtained in the presence and absence of salt will differ by less than a factor of 2. These predictions have been tested by making measurements with a number of the acids included in TABLE 1, in the presence of 0.1 M concentrations of their anions, added as the potassium salts. The results of these titrations are represented graphically in FIGURE 4. As the figure shows clearly, the difference in the pH values of the mid-points of the hydrochloric acid and benzenesulfonic acid curves are twice as great for curves obtained in the presence of salt as in its absence. The difference between the mid-points of the hydrochloric acid and naphthalenesulfonic acid curves is somewhat less than twice as great in the presence of salt as in its absence. Both differences thus agree with the predictions of the equations.

Similar experiments have also been made with Orange II at 25° in the presence of a constant concentration of its colored anion, added as the sodium salt. The affinity of this anion is so great that it combines with the protein in considerable amounts at pH values at which only small amounts of hydrogen ion are normally taken up. The resulting negative charge on the fiber is neutralized, as in all the cases covered by equations (7) and (8), by adsorption of positive ions. In the presence of salt, appreciable quantities of sodium ions, as well as hydrogen ions, are bound. Thus, equation (7) no longer gives the amount of acid taken up alone, but the sum of the amounts of acid and of salt.[20] Since the analysis of these data thus requires the development of a technique for the accurate estimation of this sum, the data obtained with Orange II at a constant concentration of its anion are not included in the present paper. The great steepness of the 0.1 M naphthalenesulfonic curve in FIGURE 4 is due, in part, to com-

bination with small amounts of salt at high pH values. These should have been added to the amounts of acid combined in order that the resulting curve should be described by equation (7).

Effect of Temperature on Anion Affinity

It has been shown earlier[2] that a comparison of titration data obtained at two temperatures permits an estimation of the heat changes accompanying the dissociation of the protein-anion complex. Measurements at 25° C. have therefore been made with four acids of widely different affinities. The results also provide a basis of comparison with those acids which it was impractical to investigate at 0° C. because of the slowness with which they come to equilibrium with wool. Data obtained at 25° C. with one of these acids, Orange II, are included with those for the other four acids in FIGURE 5. Data for three of the acids at 0° (FIGURE 1) are included for comparisons. As FIGURE 5 makes clear, the effect of temperature on the position with respect to pH of the curves obtained with picric acid (0.34 pH unit) is considerable, greater than its effect on the curves for naphthalenesulfonic acid (0.21 unit), which in turn show a greater effect than the curves for hydrochloric acid (0.16 unit). No estimate is attempted of the effect of temperature on the position of curves for flavianic acid, since the data obtained at 0° C. with this acid probably did not represent a final equilibrium.

The curve at 25° C. for flavianic acid requires further explanation. When plotted, as in FIGURE 5, as *milli-equivalents* per gram against pH, the data run nearly parallel to those for picric acid at the same temperature up to values of acid combined of at least 0.65 milliequivalents per gram. This furnishes a very strong indication that, in this range of pH values, the flavianate ions combining with the wool are predominantly doubly ionized, and each effectively neutralizes two of the positively charged basic groups in the proteins. At lower pH values, where the data are represented by a broken line, the proportion of the total flavianic acid which is doubly dissociated is so small that a part of the ions which combine with the fiber are monovalent. In the most concentrated solutions combination with the fiber takes place almost entirely with the monovalent ion, so that a plateau at 0.82 *millimole* rather than at 0.82 *milli-equivalent* would presumably be found. Since the concentration of flavianic acid is extremely low in the pH range in which the ions are predominantly divalent, it is clear that the affinity of the doubly charged anion is considerably higher than that of the monovalent ion. Trinitroresor-
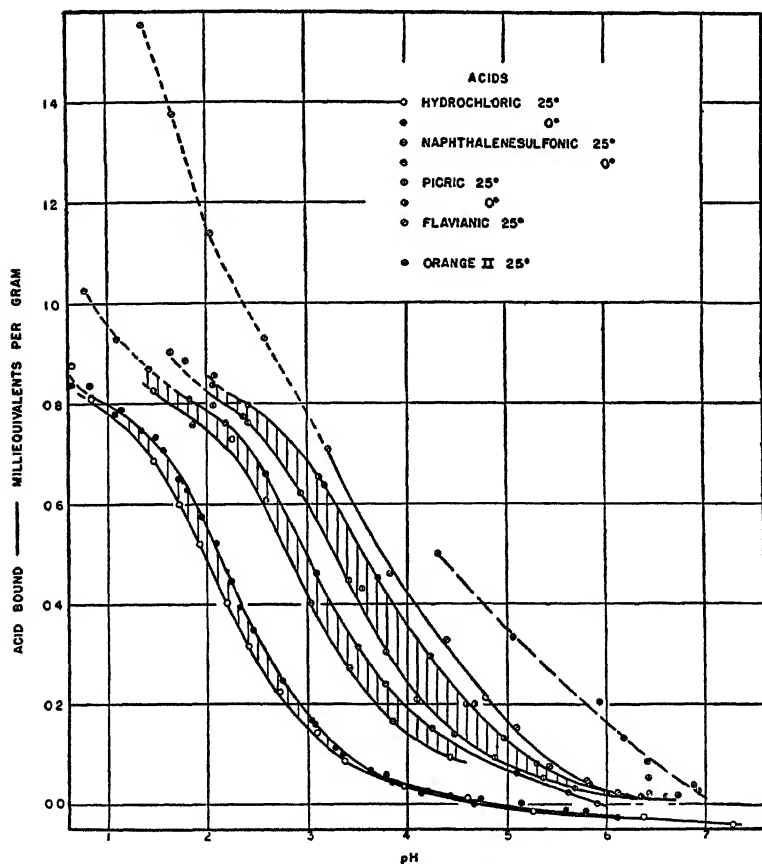
FIGURE 5. Comparison of the titration curves of wool with various acids at 0° C. and 25° C.

cinol also combines with wool as a dibasic acid when present in great dilution (to avoid complicating FIGURE 1, only the part of the data in which combination is predominantly with the monovalent ion were represented). Similar effects appear in Speakman's[9] data obtained with sulfuric acid.

The affinities at 25° C. of the acids represented in FIGURE 5, calculated by the first of the two methods described in Section 2, are given in the first part of TABLE 2. The values of K_A' listed in column 4 were obtained graphically from the curve for 25° C. in FIGURE 2. Thus the effect of temperature on K_H' has been taken into account. Owing to the inclusion of data for Orange II, the differences of affinity com-

prised in this table cover almost eight times as wide a range as those listed in TABLE 1.

The sixth column of the table gives the ratio of the protein-anion dissociation constants at 25° and 0°. From this ratio, and the van't Hoff equation, the average heat of dissociation of anions in this range of temperature has been calculated. The values so obtained are given in the last column. The heats of dissociation rise from some 2000 calories in the case of chloride ion to nearly seven thousand calories in the case of picrate. As should be expected, more energy is required to dissociate the more tightly bound ions.[25]

In order to evaluate the effect of temperature on the affinity of Orange II, the amounts of this acid combined by wool were also measured at 50° C., after allowing a period of 5 days for the attainment of final equilibrium.[26]   The results, together with data for hydrochloric acid at the same temperature, and for Orange II at 25° C., are represented graphically in FIGURE 6. The measurements represented by crosses were obtained at 50° C. by determining the increase in dry weight of the fibers as the result of combination with dye. The small discrepancy between the two methods of measuring combination (larger the more acid the solution) may be accounted for by the fact that small amounts of amide nitrogen are split off, as a result of the prolonged exposure to dye at this high temperature, and affect the apparent dye uptake estimated from the difference in the acid teter of the dye bath brought about by the wool. The amounts of wool dissolved by comparable exposure to the same concentrations of hydrochloric acid are considerably smaller.

A comparison of the titration-curve obtained with Orange II at 50° C. with the portion of the curve obtained at 25° C. strongly suggests that the data for the lower temperature do not represent a final equilibrium state. This was clearly the case when higher concentrations of dye were used,—combination of dye in 35 days at 25° C. in no case exceeded 0.63 millimole, even when solutions more acid than those used at 50° C., were employed. The affinity of the dye anion at 25° C. (listed in the first part of TABLE 2) must therefore be regarded as minimal. Thus, the heat of dissociation of the dye anions is even larger than the value in the table (12,500 calories),

    [25] The value of KH' at 25° was chosen on the basis of the earlier analysis of the effect of temperature on the combination of wool with hydrochloric acid,[2] in which the heat of transfer of the ions from the solution to the fiber was neglected.   The heats of dissociation given in the table may therefore be subject to a small correction.   Since this correction would be practically invariant, its application would not affect the relative order of the heats of dissociation, nor the great differences between them

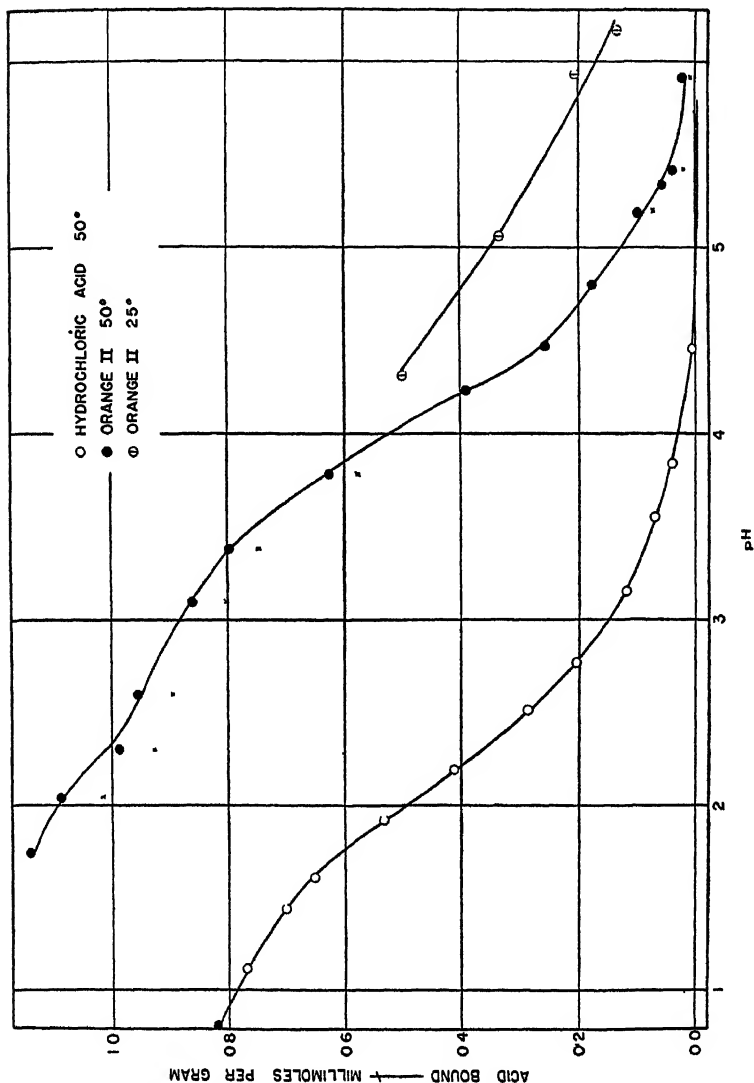   [26] Ender, W., & Müller, A. Melliand Textilberichte 18: 633. 1937.

FIGURE 6.   Combination of wool protein with the azo-dye acid Orange II, and with hydrochloric acid as a function of pH at 50° C.   The significance of the crosses is explained in the text.

For comparison, data obtained with Orange II at 25° C. are included.

TABLE 2

CALCULATION OF ANION AFFINITY CONSTANTS AND HEATS OF DISSOCIATION AT 25° AND 50° C.

Acid	pH of Midpoint	$-\log K_A'$ from eq. 8''	K_A'	Affinity relative to aff. of chloride	$\left(\dfrac{K_A'}{K_A'}\right)^{25}_{0}$	ΔH_0^{25} for anion dissociation
		25°				calories
Hydrochloric	2.16	0.31	0.49		1.40	2170
Naphthalene-β-sulfonic	3.03	1.97	.0107	46	1.81	3850
Picric	3.52	2.89	.00129	380	2.80	6670
Flavianic	4.07	3.77**	.000170**	2880**		
Orange II*	4.76	4.68†	.000021†	23400†		
		50°			$\left(\dfrac{K_A'}{K_A'}\right)^{50}_{25}$	ΔH_{25}^{50}
Hydrochloric	2.21	0.38	0.42		0.85	−1,630
Orange II*	4.21	3.97	.000107	3,900	5.09†	12,500†

* Molecular weight 328.3

** Expressed for purposes of comparison as if the anion were monovalent. Refer to text.

† The Orange II data at 25° C. do not represent equilibrium values. Thus the value of relative affinity given for this temperature, and the value of ΔH_{25}^{50} are minimal.

and exceeds by a large factor the heats of dissociation of the other ions listed.

As in experiments with several other acids, quantities of dye considerably in excess of 0.8 millimole per g were combined in the most concentrated solutions. This excess combination is not only larger with Orange II than in the other cases described but occurs in such a way as to obliterate all suggestion of a plateau in the neighborhood of 0.82 millimole per g. It therefore seems possible that causes other than those already suggested to account for the excess combination with other acids may be involved. The molecules of many dyes are known to be aggregated in aqueous solutions to a greater or lesser extent.[27] Although no evidence suggesting that Orange II is highly aggregated has ever been put forward, the failure of its solutions to obey Beer's law suggests that some degree of aggregation occurs in all but extremely dilute solutions of this dye. Combination of the fiber, at least in part, with aggregated dye anions, might then account for the large amounts taken up, and for the anomalous relation to pH of the amounts combined at high concentrations.

COMBINATION OF ANIONS WITH A DISSOLVED PROTEIN

If it can be shown that stoichiometric combination with anions occurs in the reaction of both soluble and insoluble proteins with acids, there is a wide realm of application in biochemistry[28] and physiology[29] in which it must be taken into account. If it occurs only with wool or with the fibrous proteins alone[30] it must be regarded as a clue to the nature of the specific structural differences between soluble and insoluble proteins.

In order to determine whether the marked differences in the titra-

[27] Valko, E. Trans. Faraday Soc. 31: 230. 1935.

[28] Thus, Donnan membrane potentials found in systems of acid and protein should be smaller than those predicted by calculations neglecting anion combination. With anions of especially high affinity the potentials observed might even be reversed in sign. Discrepancies between observed Donnan potentials or ion-distributions and those predicted by equations which fail to allow for anion combination have occasionally been reported (Bigwood, E. J. Bull. soc. chim. biol. 21: 1102, 1105. 1939; Briggs, D. R. Jour. Biol. Chem. 134: 261. 1940.) Calculations of membrane potentials by Adair, G. S., & Adair, M. E. (Biochem. Jour. 28: 1230. 1934; Comptes rend. Lab. Carlsberg 22: 8. 1938; Trans. Faraday Soc. 36: 23. 1940) take combination with anions into account.

[29] Physiologists have long attributed to dissolved proteins the power to form undissociated complexes with *cations* such as calcium. This combination has been characterized in several cases by a mass-action law constant (Drinker, N., Green, A. A., & Hastings, A. B. Jour. Biol. Chem. 131: 641. 1939).

[30] The amounts of acid or base combined with silk (Gleysteen, L. F., & Harris, M. Jour. Research Nat. Bur. Standards 26: 71. 1941) at any pH depend on the concentration of anions and cations respectively in a manner entirely analogous to the similar dependence described for wool.[1]

tion curves of wool obtained with different acids are also to be found with dispersed proteins, determinations were made of titration curves of the soluble protein, egg albumin, with several of the same acids. The behavior toward hydrochloric acid and chlorides of this protein has already been contrasted with that of wool in an earlier paper.[1] The measurements were made in dilute solutions of the protein
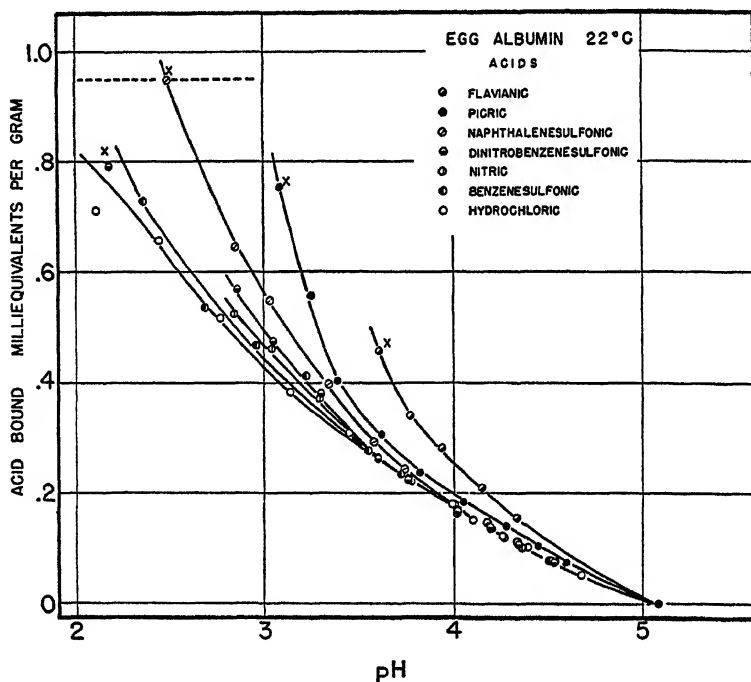

FIGURE 7.   Combination of a dissolved protein, egg albumin, with various strong acids as a function of pH.
The significance of the crosses is explained in the text.

(0.113%) in order to avoid the precipitation which would otherwise have been brought about by several of the acids used.

The results of these measurements at 22° C. are shown in FIGURE 7. The position of the broken line at the top of this figure, inserted as a reference point, shows the maximum acid-binding capacity indicated by the work of Kekwick and Cannan with hydrochloric acid.[6]   The crosses distinguish measurements in which incipient precipitation was indicated by a faint opalescence.   In every case measurements with higher concentrations of the acid produced a visible precipitate, and

have been omitted from the figure, to avoid all possibility that the effects described may have been due, in whole or in part, to the existence of two macroscopic phases.

The differences in the curves for different acids shown in FIGURE 7 evidently closely resemble the corresponding differences which have been described for wool. With the minor exception of the inversion of the relative positions of the curves for benzenesulfonic acid and nitric acid, the order of the affinities of the acids represented is the same as for wool, and the relative magnitudes of the several separations of the individual curves are qualitatively similar to the separations obtained with the fibrous protein. One difference however distinguishes the two sets of data. The positions of points along the curves for egg albumin show no tendency to remain equidistant with respect to the pH axis; neighboring curves diverge strongly as increasing amounts of acid are bound. This divergence may be regarded as a consequence of the excess of carboxyl groups over amino groups in the dissolved protein. Because of this inequality only a part of the carboxyl groups are ionized in the uncombined state, and this part alone is free to accept hydrogen ions from the acids. The complete titration curve of all the carboxyl groups of the protein extends beyond the point of zero combination with acid (near 5.0) and can only be obtained by titration with base as well as with acid. The part of the carboxyl curve which is obtained by combination with base in the absence of salt is not governed by the affinity properties of anions, since cations rather than anions are present when it is obtained. Thus, unless salt is present, the influence of anions disappears entirely at or near the point of zero combination with acid. This circumstance tends to obscure the simplicity of the relation between the affinity of the anion and the pH coordinates of acid titration curves. This simplicity will manifest itself only when the acid combination curve coincides with the titration curve of the carboxyl groups as a whole, as it does in the case of wool. With salt present, a secondary effect of the anion beyond the isoionic point may be observed, as the data of Briggs[31] obtained on serum albumen with metaphosphate plainly show.

Since the curves in FIGURE 7 converge sharply, it is difficult to compare the absolute magnitudes of the differences in affinity of the several anions with the corresponding differences found with wool. Obviously the comparison should be made between acid solutions which are as far as possible from the isoionic point in order to

[31] Briggs, D. R. Jour. Biol. Chem. **134**: 261. 1940.

minimize the effects of the convergence. The data do not extend to very acid solutions, but at the extreme of the range represented the differences in the pH coordinates of the curves approach values only slightly smaller than the differences between curves for the same acids in FIGURE 1.

Results of the kind described here for both proteins may furnish a basis for explanations of the many reports of ion specificity in the literature of protein and enzyme chemistry. Attempts have been made to deal with some of these specific properties, such as differences in salting-out effectiveness, by treatment of the electrostatic interaction between ions having different sizes, shapes, and electric moments. Many phenomena of specificity remain, however, for which plausible explanations based on simple electrostatic considerations alone are not easily obtained. Among these are the numerous instances reported of differences in the electrophoretic mobility of various proteins when different anions or different concentrations of anions are present in buffers. Such differences have been the subject of recent detailed investigations by Davis and Cohn,[32] Gorin, Abramson and Moyer,[33] Tiselius and Svensson,[34] and Sookne and Harris.[35] A number of these investigators[33, 34, 35] concluded that the differences found were due at least in part to combination of the protein with ions other than hydrogen or hydroxyl ions. An especially striking instance of anion specificity in the phenomenon of protein denaturation by guanidine salts has been reported by Greenstein.[36] Here certain salts of guanidine such as its halides are effective in low concentrations while others, such as the sulfate, are ineffective in all concentrations tried. Another example, large differences in the absolute velocity of the denaturation of dissolved crystalline pepsin in the presence of various anions, has been described by the present author. The denaturation reaction proceeds at a rate which is inversely proportional to the 5th power of the hydrogen ion activity, regardless of what buffer-acid is used to control the pH, but the rates of the reaction are widely different with each of the several acids used.[37] The shifts in the experimental curves relating any of the

[32] Davis, B. D., & Cohn, E. J.   Jour. Am. Chem. Soc. 61: 2092.   1939.
[33] Gorin, M. H., Abramson, H. A., & Moyer, L. S.   Jour Am. Chem. Soc. 62: 1643. 1940.
[34] Tiselius, A., & Svenson, H.   Trans. Faraday Soc. 36: 16.   1940.
[35] Sookne, A., & Harris, M.   Jour. Research Nat. Bur. Standards 23: 299.   1939. RP1234; 25: 47.   1939.   RP1318.
[36] Greenstein, J. P.   Jour Biol. Chem. 130: 519.   1939.
[37] Steinhardt, J.   Kg. Danske Vidensk. Selskab. Math.-Fysisk Medd. 14: No. 11. 1–53.   1937.

magnitudes cited to pH could be understood if it were shown that they were manifestations of a corresponding shift in the titration curve of the protein involved in the presence of different anions.

It has been earlier shown that the portions of the titration curves of egg albumin and of wool which are determined by the dissociation of hydrogen ions from carboxyl groups differ widely when the titrations are carried out in the absence of salt, but become closely similar when high concentrations of salt are present.[1] The effect of the presence of salt in concentrations up to 0.5 M is thus much larger with wool than with the dissolved protein. Since the theory of stoichiometric reversible anion combination was formulated to account for the large effects on wool, the application of this theory to the much smaller salt effects described by Cannan for egg albumin[38] seemed at first uncalled for. It appeared that the combination with anions might be limited to proteins present in a separate phase, possibly as a consequence of electrostatic restrictions against the transfer between phases of hydrogen ions alone. The experiments just described, however, suggest that the difference between the titration curves of dissolved and undissolved proteins in the absence of salt[1] cannot be due to the absence of any tendency for anions to combine with the dissolved protein. The contradiction between the results of the two kinds of experiment is only apparent, since the existence of a stoichiometric salt effect is not an inevitable consequence of every situation to which equation (1) may be applied. Such salt effects may be expected only when $K_A' \gg a_A \gg K_A$. If K_A' and K_A are both very large, the bracketed term in the denominator of equation (1) approaches a constant value, nearly independent of a_A. If, on the other hand, both dissociation constants are very small, the bracketed term approaches unity and is again nearly independent of a_A. In either case, corresponding to extremely low affinity or extremely high affinity of anions for the protein, the availability of hydrogen ions rather than of anions becomes the factor which limits the extent of combination with acid. Thus, the titration curves obtained in the absence of salt will not differ greatly from those obtained at low constant anion concentrations.

There are numerous indications that dissolved proteins do combine with such commonly employed anions as chloride, acetate, and phosphate but the extent of this combination does not appear to be comparable to the extent of combination of hydrogen ions in acid solutions.[28a, 28c, 33, 34] Thus, the first of the two alternatives (K_A' and K_A

[38] Cannan, R. K. Cold Spring Harbor Symp. Quant. Biol. 6: 1. 1939.

large) probably applies. The difference in the absolute values of these constants characterizing the soluble and insoluble proteins may be related to their respective states of dispersion or to the structural factors underlying them.

It should be noted in FIGURE 7 that precipitation occurred at concentrations which were in the inverse relative order of the affinities. Picrate and flavianate are ions of well known high precipitating efficiency for proteins but many others, such as tannate and metaphosphate, are equally well known. The latter in particular has been extensively investigated in recent years.[39, 40, 41, 31]   Perlmann and Herrmann have reported[41] that the metaphosphate combined is equivalent to the number of basic groups in the protein. In addition, Briggs has shown that the presence of metaphosphate results in a shift in the hydrochloric acid,—sodium hydroxide titration curve of serum albumin to a range of pH values much above that which it normally occupies. This behavior parallels the results with the other acids reported here, but the explanation which Briggs proposed is no more than formal if applied to the wide range of results with both egg albumin and wool in the present paper.

It would be of interest to determine to what extent protein precipitation by familiar specific precipitants such as picrate, tannate, and metaphosphate, and the salting out of proteins with various sulfate salts may all be regarded as manifestations of the same general protein-anion equilibrium. When the affinity is so great that the presence of only small amounts in excess of stoichiometric equivalence suffices to cause combination and precipitation, the anion or acid is naturally described as a specific protein precipitant. When the affinity is low (as in the case of sulfate) and large excesses are required, the combination and precipitation may have many of the characteristics of salting-out. With extremely low affinities, such as that of chloride, salting-out is rarely practicable.

RELATION OF ANION AFFINITY TO SIZE, STRUCTURE, AND BASICITY OF IONS

If the phenomena reported in the present and in preceding papers are to be ascribed to the formation of partially dissociated stoichiometric protein-anion combinations, it is desirable to inquire into the nature of the forces which confer so wide a range of affinities for

[39] Schofield, R. K.   Trans. Faraday Soc. 31: 390.   1935.
[40] Samuel, L. W., & Schofield, R. K.   Trans. Faraday Soc. 32: 760.   1936.
[41] Perlmann, G., & Herrmann, H.   Biochem. Jour. 32: 926.   1930.

protein to different ions. The existence in solution of ion combinations which are only partially dissociated is now well established, partially as the result of the work of Kraus and his collaborators in solutions of low dielectric constant,[42] and partially as the result of the theoretical treatment of concentrated aqueous solutions of electrolytes by Bjerrum and others. Certain partially dissociated inorganic salts (acidates, in the terminology of Bjerrum) have long been familiar, and the fact has also long been known that many substances formed by the reaction of an acid with a base, such as certain halides of silver, cadmium, and mercury, crystallize in lattices which are only partially ionic. The work of Cannan and Kibrick[13] and of Greenwald[14, 15] has shown that even in dilute aqueous solutions divalent cations often form partially dissociated phosphates and partially dissociated complexes with the anions of carboxylic acids.

Such combinations, including those considered in the present paper, may be formed in two possible ways, involving either directed or undirected valence forces. The first alternative involves the formation of a bond having a partially covalent nature, or a hydrogen bond; the second comprises various kinds of polar forces, electrostatic (coulombic) in nature or otherwise, and includes van der Waals forces. If the first type is involved the principle relation found among the anions should be an inverse one of affinity to the intrinsic strength of the corresponding acid.[43] The relations found in cases to which the second alternative applies would be expected to be more complex.

Since most of the acids included in the present research are of approximately equal strength in water, but differ widely in affinity, the first alternative seems to have little in its favor. However, a part of the total binding energy between anion and protein might still be due to the formation of a directed bond, especially since the intrinsic strengths of the various acids, as distinguished from their apparent strengths in water, are not all equal.

In an effort to appraise more closely the influence of the basicity of the anion on the affinity, comparisons have been made of the results obtained with pairs of comparable acids, such as benzenesulfonic and benzoic acids, which differ widely in strength (FIGURE 8). The data shown are limited to ranges of pH in which the concentration of the undissociated form of the weak acid does not exceed 0.1 M in

---

[42] **Kraus, C. A.** Science, **90**: 281. 1939.
[43] Here the possibility of the formation of a bond between the protein and the charged group of the anion (when the anion is polyatomic) is alone under consideration.
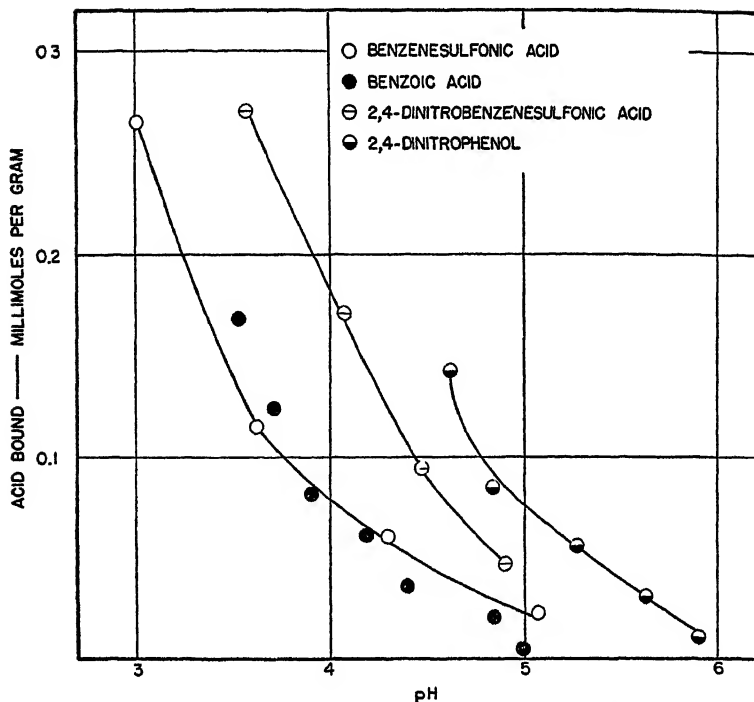
Figure 8. Comparison of portions of the acid titration curves of wool protein obtained with pairs of homologous acids which differ widely in strength.

order to avoid effects, due to the presence of larger amounts, which have been described elsewhere.[11] There is very little difference between the amounts of strong and of weak acid combined at a given pH in the case of the benzenesulfonic-benzoic acid comparison. On the other hand, dinitrophenolate ion appears to have a definitely higher affinity for wool than the anion of its totally dissociated analogue, dinitrobenzenesulfonic acid. The opposite situation is found in experiments (still in progress) with a series of aliphatic carboxylic acids. These have even lower affinities for wool than has hydrochloric acid; in addition, the weaker acids have lower affinities than the relatively strong members of the series. The affinities of all of them are increased by the introduction of a halogen or of a hydroxyl group, although the derivatives thus formed are considerably stronger acids. Since there is no consistency in these results it seems likely that the differences found are due to specific structural effects rather than to differences in acid strength.

An indication of the nature of the *undirected* forces to which the principal recourse must therefore be had is to be found in the fairly regular correspondence between the order of affinities and of molecular weights, which is evident in TABLE 1. Such a dependence on the mass of the anion is consistent with the Fajans rule which deals with the tendency of an ionic bond to dissociate. However, a dependence on mass, or functions of ion size related to it, is also generally characteristic of similar phenomena which depend at least in part on van der Waal's forces, or adsorption.[44, 45]

The several secondary exceptions to a strict correspondence between the relative masses of the ions and their affinities, provide further information as to the relation between structure and affinity. Thus the aliphatic carboxylates already mentioned, do not fall into a molecular weight order; likewise, ethylsulfate ion has as low an affinity as the much smaller chloride ion, and is lower in relative affinity than the smaller bromide and nitrate ions. The anion of trinitroresorcinol is slightly lower in affinity than picrate although it is slightly heavier. Naphthalenesulfonate is higher in affinity than the heavier dinitrobenzenesulfonate ion.

A possible way of resolving these discrepancies without introducing a host of special assumptions suggests itself if due regard is had for the shapes of the various ions, and the character of the surfaces which they present to the protein. Two-dimensional ions (those derived from simple benzenoid compounds) possess higher affinities than either compact three-dimensional or chain-like particles of the same mass. The nature of the substituents added appears to be of less importance than their mass as long as there is no steric hindrance to their lying in the plane of the ring. Such a hindrance appears to operate in the case of the monovalent trinitroresorcinol ion which has an appreciably lower affinity than picrate. It appears doubtful, from the study of models, whether the undissociated hydroxyl group in this ion can lie in the plane of the ring. When the flat configuration is imposed by the existence of conjugated double bonds, as in naphthalenesulfonic acid, the affinity relative to a disubstituted benzenoid compound of the same mass gains accordingly. Wide differences between the degree of substantivity of *o*- and *m*- substituted benzidine azo-dyes cited by Robinson and Mills[45] furnish extreme examples which are entirely consistent with the present suggestion.

[44] The Langmuir adsorption law leads to the same equation as the law of mass action in the derivations of the equations describing the combination of acid.[1]

[45] Hendricks, J. B. Jour. Phys. Chem. 45: 65. 1940.

[45] Robinson, C., & Mills, H. A. T. Proc. Roy. Soc. London, A131: 576. 1931.

The fact that ion-flatness appears to favor affinity for protein need not be interpreted as an exclusive dependence of affinity on the size of the surface presented by an ion. Saturated aliphatic hydrocarbon chains in an ion probably do not exercise any considerable attractive forces on the fragments of hydrocarbon chains which exist on the surface of the protein. A close approach between them is obstructed by the outer shell of hydrogen atoms common to both. Thus, the greater affinity of ions composed of aromatic hydrocarbons may be due to the absence of this obstructive shell on two of their largest surfaces. If, in addition, electro-negative substituents, such as chlorine or the oxygen of nitro groups, are presented by the ion to the protein, strong attractive forces may arise, replacing the relative indifference which determines the behavior toward protein of the unsubstituted hydrocarbon. Experiments designed to test these views, and the possibility of applying them in the deliberate control of affinity, are at present in progress.

ACID-BASE EQUILIBRIUM IN SOLUTIONS
OF AMPHOLYTES

By John G. Kirkwood

From Cornell University

Important information concerning the structure of the amino acids and proteins is provided by a study of acid-base equilibria in solutions of these ampholytes. In fact, one of the early arguments in favor of the dipolar ion hypothesis[1] was based upon the striking discrepancy between the values of the dissociation constants of the amino acids and those of aliphatic carboxylic acids. The acidic group of the neutral ampholyte was concluded to be $NH_3{}^+$ rather than COOH, in agreement with a dipolar ionic structure $NH_3{}^+RCOO^-$. On the basis of Bjerrum's theory of electrostatic interaction in the dissociation of polybasic acids, it is further to be expected that the negative group COO^- would have a marked influence on the acidic dissociation of the $NH_3{}^+$ group, depending upon the distance of separation of the charged groups. From a comparison of the dissociation constants of the dipolar ion $NH_3{}^+RCOO^-$ and the ion of its ester salt $NH_3{}^+RCOOCH_3$, the charge separation in the dipolar ion may be calculated. Neuberger[2] has calculated the charge separation for several aliphatic amino acids on the basis of the Bjerrum theory. The distances so obtained were considerably too small. More recently Westheimer and Shookoff[3] using a more refined electrostatic theory, have computed charge separations which agree very satisfactorily with structural estimates from accepted interatomic distances and bond angles. Their calculations thus complete the argument for dipolar ionic structure based upon the magnitude of dissociation constants. It also is possible to extend the electrostatic theory to an ampholyte with an arbitrary number of acidic groups, $NH_3{}^+$ and COOH, and thus to provide a semi-quantitative theory of acid-base equilibria in solutions of proteins.

The ionization equilibria of a simple ampholyte HZ with positive ion H_2Z^+ and negative ion Z^- are described by the equations;
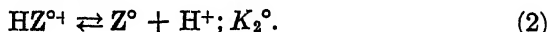
$$H_2Z^+ \rightleftarrows HZ + H^+; K_1$$

$$HZ \rightleftarrows Z^- + H^+; K_2. \tag{1}$$
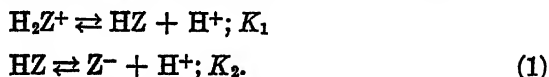
[1] Bjerrum, N. Zeit Physik. Chem. **106**: 219. 1923.
[2] Neuberger, A. Proc. Roy. Soc. London **A158**: 68. 1937.
[3] Westheimer, F. H., & Shookoff, M. W. Jour. Am. Chem. Soc. **61**: 555. 1939.
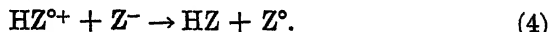
Let $HZ^{\circ+}$ be a monobasic acid with a basic part Z° differing in structure from the ion Z^- only by the absence of the negative charge. Its ionization equilibrium is described by

$$HZ^{\circ+} \rightleftarrows Z^\circ + H^+; K_2^\circ. \tag{2}$$

From thermodynamics we may write

$$RT \log K_2^\circ/K_2 = -\Delta F^\circ \tag{3}$$

where ΔF° is the standard free energy increment in the reaction,

$$HZ^{\circ+} + Z^- \rightarrow HZ + Z^\circ. \tag{4}$$

From a molecular point of view, $\Delta F^\circ/N$ is equal to average work expended in the transport of a proton from the molecule Z° to the ion Z^-. Since Z° and Z^- contain the same basic group, the effect of short range exchange forces bonding the protons in the acids $HZ^{\circ+}$ and HZ may be expected to cancel in the transfer of the proton between them, leaving the electrostatic interaction between the proton and the negative charge of Z^- as the dominant contribution to ΔF°. If R is the distance between proton and the negative charge in the molecule HZ, ΔF° then has the form

$$-\Delta F^\circ = \frac{Ne^2}{\epsilon' R} \tag{5}$$

where ϵ' is an effective dielectric constant. In the Bjerrum theory ϵ' was assumed identical with the macroscopic dielectric constant ϵ of the solvent. In the Kirkwood-Westheimer theory,[4] the molecule HZ is treated as a cavity of low dielectric constant rather than as a structureless system of two point charges in the solvent continuum. Using classical electrostatic theory, these authors have calculated ϵ' for molecules of spherical and ellipsoidal shape. From equations (3) and (5) the quantity Δp_K, equal to $\log_{10} K_2^\circ/K_2$, may be computed from the formula,

$$\Delta p_K = \frac{e^2}{2.303\epsilon' RkT} \tag{6}$$

where ϵ' may be obtained from the tables of Kirkwood and Westheimer. In their spherical model, the molecules HZ and $HZ^{\circ+}$ are regarded as spheres of radius b, each containing a proton at a distance r from their centers. The molecule HZ is assumed to contain a nega-

[4] Westheimer, F. H., & Kirkwood, J. G. Jour. Chem. Phys. 6: 506. 1938; 6: 513. 1938.

tive charge e, in excess of HZ^{o+}, situated at a distance r from the center on a vector making an angle ϑ with the vector to the proton. The values of ϵ' in the solvent water at 25° are presented in TABLE 1. They were calculated with an internal molecular dielectric constant equal to 2.00.

TABLE 1

$(x = r/b)$

VALUES OF $\sqrt{x}\,\epsilon'$

$\sqrt{x}/\cos\vartheta$	-1	$-\frac{1}{2}$	0
0 10	0 25	0 24	0 23
.20	.64	.60	55
30	1 31	1 16	1 02
.40	2 45	2 11	1 76
.50	4 59	3 82	3 02
.60	8 79	7 15	5 42
.70	17 6	14 2	10 6
80	37 5	30 9	23 3
90	79 2	68 9	56 4
1 00	127 0	116 3	103 5

For the calculation of R from experimental values of R, it is convenient to write equation (6) in the form,

$$x\epsilon'(x) = \frac{e^2}{4.606\,bkT\,\Delta p_K \sin\vartheta/2} \tag{7}$$

$$= \left(\frac{3v}{4\pi N}\right)^{\frac{1}{3}}$$

$$R = 2bx \sin\vartheta/2.$$

where v is the molar volume of the acid HZ. With an appropriate structural assignment of ϑ, the function $\sqrt{x}\,\epsilon'\,(x)$ may be computed and the corresponding value of x obtained from TABLE 1. From the latter value the intercharge distance R is calculated from the third of equation (7).
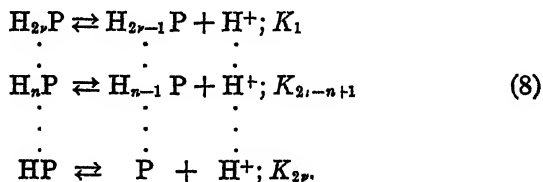
If the acid HZ is identified with the dipolar ion $NH_3^+RCOO^-$ and the acid HZ^{o+} with the ester salt ion $NH_3^+RCOOCH_3$, the theory may be used to calculate the distance between the charged group of the dipolar ion. Such calculations have been made by Westheimer and Shookhoff, using the spherical model for glycine and alanine and the ellipsoidal model for amino acids and peptides of longer chain length. Their results are presented in TABLE 2, together with the dis-

TABLE 2

CHARGE SEPARATION IN DIPOLAR IONS FROM Δp_K VALUES

	Δp_K	R	R_F	R_B
Glycine	2.02	4.05	3 56	1.53
Alanine	2 07	3.85	3 56	1 50
β-Alanine	1.06	5.15	4.19	2.92
γ-Aminobutyric acid	.72	6.10	4 72	4.31
δ-Amino valeric acid	.62	6.55	5 19	5.00
ε-Amino caproic acid	.38	7.85	5.63	8.16
Glycylglycine	.56	6.50	5.17	5 54

tance R_F computed on the basis of free rotation and R_B on the simple Bjerrum theory in which ϵ' is identified with the solvent dielectric constant. The distance 4.05 Å obtained for glycine agrees moderately well with the structural value 3.17 and the salting-in value 3.30. The mean charge separations in the amino acids of longer chain length, as those computed from salting-in and from the dielectric constant increment, do not differ greatly from the free rotation values. Finally, we may say that the dissociation constant data on amino acids and their ester salts are entirely in harmony with the dipolar ion hypothesis.

It is possible to extend the ideas which have just been discussed to ampholytes containing an arbitrary number of acidic and basic groups. We shall begin with some preliminary remarks on the thermodynamic aspects of the dissociation of an ampholyte $H_{2\nu}P$ capable of forming a series of ions $H_{2\nu}P \ldots H_nP \ldots P$ with respective charges $n-\nu$. The dissociation equilibria are described by the equations:

$$H_{2\nu}P \rightleftarrows H_{2\nu-1}P + H^+; K_1$$

$$\vdots \quad \vdots \quad \vdots$$

$$H_nP \rightleftarrows H_{n-1}P + H^+; K_{2\nu-n+1} \tag{8}$$

$$\vdots \quad \vdots \quad \vdots$$

$$HP \rightleftarrows P + H^+; K_{2\nu}$$

If x is the hydrogen ion activity $C_P{}^\circ$ the bulk ampholyte concentration, and C_{H_nP} the concentration of the ion H_nP, we may write,

$$C_{H_nP} = x\gamma_{n-1} C_{H_{n-1}P}/\gamma_n K_{2\nu-n+1}$$

$$C_P{}^\circ = \sum_{n=0}^{2\nu} C_{H_nP} \tag{9}$$

where γ_n is the activity coefficient of the ion H_nP. The difference equations (9) for the concentrations C_{H_nP} have the solutions,

$$C_{H_nP} = C_P{}^\circ (\prod_{s=1}^{n} K_{2\nu-s+1}{}^{-1})x^n/\gamma_n G(x)$$

$$G(x) = \sum_{k=0}^{\hat{}} (\prod_{s=1}^{\cdot} K_{2\nu-s+1}{}^{-1})\gamma_k{}^{-1}x_k. \tag{10}$$

Thus the fraction f_n of the ampholyte existing in the form of the ion H_nP is

$$f_n = (\prod_{s=1}^{n} K_{2\nu-s+1}{}^{-1})x^n/\gamma_n G(x). \tag{11}$$

The mean charge $\bar{Z}$ of the ampholyte is evidently given by

$$\bar{Z} = \sum_{n=0}^{2\nu} (n-\nu)f_n = \frac{d \log G}{d \log X} - \nu. \tag{12}$$

The isoelectric point corresponding to vanishing average charge is obtained by solving the algebraic equation

$$\frac{d \log G}{d \log X} - \nu = 0 \tag{13}$$

for x, equal to 10^{-pH}. The activity coefficient of the ampholyte, regarded as the neutral species H_nP is $f_\nu\gamma_\nu$, which may be written

$$\gamma = \gamma_\nu/(1 + \gamma_\nu\varphi(x)) \tag{14}$$

$$\varphi(x) = \sum_{n=0}^{\nu-1} (\prod_{s=n+1}^{\nu} K_{2\nu-s+1})\gamma_n{}^{-1}x^{n-\nu} + \sum_{n=\nu+1}^{2\nu} (\prod_{s=\nu+1}^{2\nu} K_{2\nu-s+1}{}^{-1})\gamma_n{}^{-1}x^{n-\nu}.$$

In an aqueous solution containing the ampholyte and a strong acid at a concentration C_A, electrical neutrality requires

$$\sum_{n=0}^{2\nu} (n-\nu)f_n C_P{}^\circ + {}_{\cdot}C_H{}^+ - C_A - C_{OH-} = 0. \tag{15}$$

The bound hydrogen ion per mole of ampholyte, y, is defined as follows

$$C_H{}^+ = X/\gamma_H{}^+ = C_A + C_{OH-} - yC_P{}^\circ \tag{16}$$

and the bound hydrogen ion becomes equal to the average charge, $\bar{Z}$, or

$$y = \frac{d \log G}{d \log X} - \nu. \tag{17}$$

All pertinent information concerning the acid-base equilibria of the ampholyte may thus be calculated from the function $G(x)$

$$G(x) = \sum_{n=0}^{2\nu} B_n x^n \tag{18}$$

$$B_n = \gamma_n^{-1} \prod_{s=0}^{n} K_{2\nu-s+1}^{-1}.$$

We shall now undertake the development of an approximate statistical theory of the function $G(x)$ for a large spherical ampholyte of radius R, containing ν acidic groups COOH and ν acidic group NH_3^+, following the Linderstrøm-Lang[5] treatment. Let us suppose that K_1° and K_2° are the respective dissociation constants of an isolated COOH group and an isolated NH_3^+ group in the absence of electrostatic interaction between the protons of the ampholyte ion. In the ion H_nP, the n protons can assume a number of configurations corresponding to the different ways in which the ν basic sites COO^- and the ν basic sites NH_2 may be assigned to them. If W_c is the local free energy of the configuration C, the configurational part of the chemical potential μ_n° of the ion H_nP may be computed from the partition function,

$$e^{-[\mu_n{}^\circ - n\mu_H{}^{+\circ}]/kT} = \sum_c e^{-W_c/kT}$$

$$W_c = n_1 kT \log K_1^\circ + n_2 kT \log K_2^\circ + V_c \tag{19}$$

where n_1 is the number of protons occupying COO^- sites, n_2 the number occupying NH_2 sites, and V_c is the mutual electrostatic energy of the n protons and of each proton with the negative charges of the COO^- groups other than that occupied by it. The calculation of V_c demands a specification of the location of the basic groups on the surface of the molecule and use of the Kirkwood-Westheimer theory. We shall assume that the basic sites are randomly distributed on the surface of the sphere R. On the Kirkwood-Westheimer theory, the mean electrostatic interaction of any pair of charges averaged over all points on the surface of the sphere is simply $e_1 e_2/2\epsilon R$ where ϵ is the dielectric constant of the solvent. The average electrostatic energy may therefore be roughly approximated by

$$\bar{V} = [n(n-1) - 2n_1(\nu-1) - 2n_2\nu]e^2/2\epsilon R \tag{20}$$

If fluctuations in V_c are neglected in the calculation of the partition function (19), we may now write

⁵ Linderstrøm-Lang, K. Compt. rend. trav. lab. Carlsberg **15** (7): 1923-25

$$e^{-[\mu_n{}^\circ - n\mu_H{}^{+\circ} - \mu_0{}^\circ]/kT} = e^{\nu 2\alpha - (n-\nu)2\alpha A_n}$$

$$\alpha = e^2/2\varepsilon RkT$$

$$A_n = \sum_{\substack{n_1,\,n_2 \\ = 0 \\ n_1+n_2=n}}^{n} \binom{\nu}{n_1}\binom{\nu}{n_2} \lambda_1{}^{n_1} \lambda_2{}^{n_2} \tag{21}$$

$$\lambda_1 = (K_1{}^\circ e^\alpha)^{-1} \qquad \lambda_2 = (K_2{}^\circ e^{-\alpha})^{-1}$$

since $\binom{\nu}{n_1}\binom{\nu}{n_2}$ is the number of configurations corresponding to n_1 protons on COO^- sites and n_2 on NH_2 sites. By thermodynamics, we have

$$kT \log K_{2\nu-n+1} = \mu_n{}^\circ - \mu_{n-1}{}^\circ - \mu_{H^+}{}^\circ \tag{22}$$

and from equations (21) and (22) we may write

$$K_{2\nu-n+1} = \frac{A_{n-1}}{A_n} e^{(2n-2\nu-1)\alpha} \tag{23}$$

Thus the coefficient B_n in the function $G(x)$ becomes

$$B_n = \gamma_n{}^{-1} \prod_{s=1}^{n} K_{2\nu-s+1}{}^{-1} = e^{\nu 2\alpha - (n-\nu)2\alpha - \log \nu_n} A_n \tag{24}$$

In all calculations the factor $e^{-\nu^2\alpha}$ is of no importance, so that it is sufficient to write

$$G(x) = \sum_{n=0}^{2\nu} e^{-(n-\nu)2\alpha - \log \gamma_n} A_n(\lambda_1, \lambda_2) x^n \tag{25}$$

$$A_n(\lambda_1, \lambda_2) = \text{Coefficient of } t^n \text{ in } (1+\lambda_1 t)^\nu (1+\lambda_2 t)^\nu$$

We shall not discus in detail the electrostatic effect embodied in the factor $e^{-(n-\nu)^2\alpha}$, except to remark that it favors the dipolar ion form of the ampholyte $H_\nu P$. However, due to the fact that α may be small in a protein molecule, the electrostatic interaction between the protons of the ampholyte may be dominated by the interaction with the electrolytic environment, manifesting itself in the $\log \gamma n$ term. Thus acid-base equilibrium and the isoelectric points may be expected to be sensitive to the ionic strength of the ampholyte solution. It is perhaps a matter of interest to state the approximation to $G(x)$ with the neglect of electrostatic interaction within the ampholyte molecule and with the environment.

$$G(x) = (1+x/K_1{}^\circ)^\nu (1+x/K_2{}^\circ)^\nu \tag{26}$$

By equation (12), the mean charge of the ampholyte becomes

$$\bar{Z} = \nu\left\{\frac{x}{K_1^\circ + x} + \frac{x}{K_2^\circ + x} - 1\right\} \tag{27}$$

For the isoelectric point, $\bar{Z} = o$, we have

$$x = (K_1^\circ K_2^\circ)^{1/2} \tag{28}$$

a value independent of the ampholyte's specific structure. This corresponds to a pH of approximately 6. The specific behavior of an ampholyte is thus to be attributed to electrostatic interaction between its charges and with its electrolyte environment. The theory may be extended without difficulty to an ampholyte containing acidic groups other than NH_3^+ and $COOH$, for example, NH_2^+ and SH groups.

SENSORY CUES USED BY RATTLESNAKES IN THEIR RECOGNITION OF OPHIDIAN ENEMIES*

By

Charles M. Bogert

CONTENTS

Several years ago Klauber (1927: 13) reported that rattlesnakes seemed to recognize the king snake as an enemy. He noted that small captive rattlesnakes cornered by king snakes reacted in a peculiar manner, placing the head on the ground and raising the body in a vertical loop. Klauber believed this action hindered the king snake in securing a hold on the head.

More recently Cowles (1938: 13) has reported extensive observations on the behavior of two species of rattlesnakes (*Crotalus viridis oreganus* and *C. cerastes*) in the presence of king snakes. He has described the unusual defense posture assumed by the rattlesnakes, pointing out that the loop of the body is used to repel an enemy in much the same fashion "as a human being would use an elbow in striking a heavy blow." The posture assumed in the presence of the ophiophagous snakes was in contrast to the more usual defense posture of rattlesnakes in the presence of mammalian enemies, that is, with the head raised, in preparation for a strike forward. Cowles noted a certain reluctance on the part of individual king snakes (*Lampropeltis getulus californiae*) to attack rattlesnakes and he believed this to be due to the king snake's fear of being struck by the rattlesnake's raised body. He adds that "king snakes of the species under discussion seem able to identify rattlesnakes either by sight or by smell."

No attempt appears to have been made to identify the senses used by rattlesnakes in their recognition of ophidian enemies. However, Cowles asserted that the same defense posture assumed by rattlesnakes in the presence of ophidian enemies was elicited by the odor of the spotted skunk (*Spilogale*) on a person who had been skinning this mammal.

In the present investigation an effort has been made to determine: (1) what sensory cues are used by rattlesnakes in their discrimination between ophiophagous snakes and those which normally do not prey upon rattlesnakes and (2) whether rattlesnakes react only to the species of ophiophagous serpents found in their respective native habitats or to other species as well.

MATERIAL

In preparation for this investigation a collection of freshly captured rattlesnakes was assembled.[1] Most of these were sidewinders (*Cro-*

talus cerastes) secured in the vicinity of Indian Wells, Riverside
County, California, but six additional species and subspecies were
obtained from various parts of the West (California, Arizona, Utah,
and Montana).   In addition to these crotalids, several harmless
snakes were utilized in experiments.   Most of these snakes were
colubrid species, although the family Boidae was represented by two
species.

ROLE OF VISION IN ENEMY RECOGNITION

For many years it has been recognized that many species of king
snakes included in the genus *Lampropeltis* are ophiophagous.   There
are numerous accounts of captive specimens eating rattlesnakes as
well as other serpents (Blanchard 1921; Klauber 1927: 13).   More
satisfactory evidence that king snakes prey upon crotalids under field
conditions is reported by Klauber (1931: 71), who states that a re-
cently collected specimen of *Lampropeltis getulus californiae* dis-
gorged three juvenile Pacific rattlesnakes (*Crotalus viridis oreganus*),
and a second specimen found dead on the highway had eaten a Pacific
rattler about half its own length.

A method of investigating the role of vision in enemy recognition
was afforded by using a species of king snake having two pattern
phases.   Klauber's recent papers (1936, 1939) have established the
fact that pattern dimorphism in *Lampropeltis getulus californiae*
accounts for the presence of both longitudinally striped phases and
ringed pattern phases in the coastal area of southern California.   The
striped form as well as the ringed form may occur in the same brood.
On the deserts of California the ringed form of *L. g. californiae*
(formerly known as *L. g. boylii*) is found in limited numbers, usually
in the vicinity of springs but occasionally several miles from water.
Not infrequently it is taken in the same habitat with *C. cerastes*.
The striped form is known to occur on the desert, only in San
Diego County and, therefore, is rarely found within the habitat of
the arenicolous sidewinder.

If sidewinders rely upon visual cues in their recognition of king
snakes, it seemed unlikely that any fear of the striped Pacific king
snake would be exhibited.   To subject this to test an adult striped
king snake was introduced into a cage 70 by 40 centimeters, contain-
ing four adult sidewinders which measured from 40 to 60 centimeters
overall.   The sidewinders were undisturbed prior to this test and all
of them had assumed the flattened "resting coil" characteristic of the
species.   However, as the king snake crawled about the cage the

rattlesnakes reacted immediately when the king snake approached. Inflating their bodies, but holding their heads flush with the bottom of the cage, the rattlesnakes arched the middle of the trunk region and from time to time delivered a blow with the raised portion.   The king snake sought only to escape, and made no attempt to seize any of the sidewinders.

This snake was then removed and a larger specimen of the same longitudinally striped king snake was placed in the cage with the sidewinders.   This time the king snake immediately attacked a large, gravid, female sidewinder.   As the king snake approached, the sidewinder arched her body as she had previously done, but when the king snake attempted to seize her by the neck a sudden movement of the sidewinder jerked the neck free, and simultaneously she struck a fending blow with her looped body.   The king snake abandoned the fight, but crawled to a distant side of the cage where it cautiously approached a smaller sidewinder, a male.   This time the king snake seized the rattler by the head despite his attempt to defend himself. Maintaining its grip on the head, the king snake swiftly placed several coils around the trunk of the struggling sidewinder and began to engulf it.   At this juncture the experimenter interfered, tapping the king snake on the head in an effort to make it disgorge its prey. The king snake loosened its coils but continued to engulf its victim, advancing one set of maxillary teeth while it held its prey fast with the set on the opposite side.   Further efforts to discourage the king snake were abandoned, and within seven minutes the sidewinder, a specimen about 30 centimeters in length, was completely swallowed.

The presence of a striped king snake invariably elicited the same reaction in the rattlesnake.   Rattlesnakes assumed the characteristic defense posture no more readily when the ringed form was present.

Similar experiments were repeated ten times at separate intervals with four king snakes (two of each pattern phase) and nine sidewinders as well as with individual rattlers of several species (*C. s. scutulatus, C. atrox, C. v. viridis*, and *C. v. lutosus*).   It seemed doubtful whether visual stimuli played any part in the rattlesnakes' recognition of king snakes.   Further data bearing out this conclusion are cited in the following account.

ROLE OF CHEMORECEPTION IN ENEMY RECOGNITION

Various workers have reported observations which indicate that chemoreception plays an important role in the courtship, feeding, and other activities of snakes.   Baumann (1929) claimed that the male

of *Vipera aspis* employed chemical cues in following the trail of a female during the breeding season. Among other conclusions drawn from a study of the aggregation behavior of a colubrid snake, *Storeria dekayi*, Noble and Clausen (1936: 315) stated, "Species identification is accomplished chiefly through an olfactory sense, while vision enters into the response secondarily. The tongue itself plays no part in species identification." In a paper dealing with the courtship of snakes, Noble (1937) concluded, "Sex recognition is accomplished chiefly by chemical sense, both olfactory and Jacobson's organs playing a part." He showed that the body skin itself is the chief source of sexual attraction. Kahmann's (1932) experiments indicated that the organs of Jacobson, the epithelial lining of which contain typical olfactory cells whose extensions form the vomeronasal nerve, are the only receptors used by the European grass snake in the trailing of prey. Baumann (1929) at an earlier date had concluded that chemoreception played the dominant role in the location of prey by vipers. Employing an improved technique Wilde (1938) reports verification of the conclusion drawn by Kahmann concerning the role of Jacobson's organs, although Wilde was dealing with the feeding reaction of the common garter snake. He believes the determining factor in this reaction to be chemical stimulation through the organs of Jacobson. Since snakes use their tongues to pick up minute quantities of odorous materials in the air and to carry them to the organs of Jacobson, Wilde found it possible to abolish the feeding response either by severing the vomeronasal nerve or by ablation of the tongue.

The latter method adopted by Wilde was utilized in some of our experiments described below. However, before the tongue was removed from any of the rattlesnakes some preliminary tests were performed. Within two weeks of their capture thirteen *C. cerastes*, one *C. s. scutulatus*, one *C. m. molossus*, and one *C. atrox* were placed in a circular enclosure approximately two meters in diameter. At intervals of at least a day three subspecies of *Lampropeltis getulus*, i. e., *californiae*, *yumensis*, and *splendida*, respectively, were introduced into the enclosure. Usually the experimenter was forced to guide the movements of the king snakes toward the rattlesnakes. Sometimes the rattlers reacted by lowering their heads and raising the vertical loops of their trunks; at other times they merely hid their heads beneath their coils, although prodding them with a stick evoked the raised loops. (In the absence of king snakes, prodding normally brings forth the more usual defense reaction, with the head raised.) Not infrequently sidewinders fled precipitately, attaining considerable

speed with their peculiar method of locomotion (for a description of "sidewinding," see Mosauer 1930). Other species of rattlesnakes rarely attempted to flee. Repeated maneuvering of the king snakes by the experimenter eventually evoked the characteristic defense posture in all rattlesnakes except the *C. m. molossus*.

Whereas rattlesnakes normally assume an attitude of defense with the head raised in preparation for the strike when confronted by man or by domesticated animals, immediately after exposure to king snakes they react to contact or visual stimuli only by reassuming the "king-snake defense posture." After all king snakes had been removed from the enclosure, abrupt movements accompanied by the lowering of the head and the raised trunk continued to be manifested each time any of the rattlers touched one of the others. Three hours after all king snakes had been removed from the enclosure the rattlers had assumed "resting" positions, with their heads placed on the outer edge of circular coils. At this time most of them raised their heads preparing to strike when confronted by the experimenter. However, two sidewinders and the Mojave rattlesnake, *C. s. scutulatus,* still reacted to prodding with a stick by assuming the position referred to as the "king-snake defense posture." Each time a stick was applied to the head or neck of these three rattlers the reaction involved a sudden withdrawal of the head accompanied by a violent shift in the position of the entire body. Application of the stick to other parts of the body was not always sufficient to elicit this response.

These preliminary tests were made on a warm afternoon when the temperature was approximately 34° C. within the enclosure, an abandoned cistern on the ranch belonging to Mr. Melville H. Haskell, near Tucson, Arizona. Three weeks later, after the rattlesnakes and some of the king snakes had been shipped to the laboratory in New York City, the tests described above were repeated. At this time (June 10), four of the sidewinders no longer responded to king snakes and these four, together with *C. m. molossus* which still failed to respond even after it had shed its skin, were eliminated from subsequent tests. Meanwhile, additional rattlesnakes became available. These included a specimen of *Crotalus viridis lutosus* from Utah and four *C. v. viridis* from Montana. Four Pacific king snakes (two of each pattern phase) used in subsequent tests were all from the coastal region of San Diego County, California.

To repeat a fortuitous observation made in 1935, an experiment was carried out under controlled conditions in the laboratory. Two king snakes were placed in a gallon-sized glass jar where they re-

mained for several hours. Five minutes after the king snakes had been removed and taken to another room, a sidewinder was placed in the jar. The lowered head and abrupt movements of the sidewinder were taken to indicate that the reaction had been evoked. The same procedure was accompanied by similar results when four additional sidewinders and a *C. s. scutulatus* were used. Five sidewinders and the *C. v. lutosus* did not react when placed in a clean jar not previously occupied by king snakes. One additional phenomenon observed is worth noting. Whereas the crotalids used in the experiments nearly always rattled in response to the presence of the experimenter, vibration of the tail ceased abruptly when the reaction to the king snakes was evoked. This was noted both in the preliminary experiments in Arizona as well as in the tests made by placing rattlesnakes in jars which had contained king snakes. Rattlers placed in clean jars either continued to rattle and thrust their heads up the edge of the glass, or settled their bodies in resting coils.

Additional evidence that chemoreceptors are utilized by rattlesnakes in their recognition of king snakes was derived from the following experiment. Four *C. cerastes*, one *C. s. scutulatus*, and one *C. atrox* were blindfolded. This was accomplished by means of strips of adhesive tape applied to the eyes, care being taken to make certain that each eye was completely covered. Blindfolded rattlesnakes placed with king snakes assumed the characteristic defense posture as readily as five *C. cerastes*, four *C. v. viridis*, and one *C. v. lutosus* not blindfolded.

Using the same series of rattlesnakes, an attempt was made to discover the source of the stimulus. Clean sticks of pine thrust in front of rattlesnakes caused them to extend the tongue in repeated flickerings. The snakes without blindfolds usually withdrew from the stick, making an effort to escape from the cage. Snakes with blindfolds extended their tongues more frequently and usually approached the stick, particularly if it was waved gently in front of them; they retreated only if touched on the snout, but did not assume the "king-snake defense posture."

However, sticks prepared in divers ways produced different results when held in front of rattlesnakes. Three pine sticks, approximately a meter in length and three centimeters in diameter, were made ready on separate days. The end of one stick was smeared with the secretion from the cloacal glands of the king snakes; a second stick was rubbed vigorously on the dorsum of a king snake, and a third stick was rubbed on the ventral surface. In each case tests were made

immediately after sticks were prepared. The method used was the same with each rattlesnake, including the six specimens that were blindfolded. In each case a clean stick was first held immediately in front of the head of each rattlesnake before the prepared stick was used. Throughout these tests manipulations were performed as quietly as possible; that is, a minimum amount of vibration was produced in the room. Whereas rattlesnakes apparently do not hear (Manning 1923), they are extremely sensitive to vibrations received through the substratum and respond by rattling to such stimuli received through the bottoms of the cages. Rattlesnakes without blindfolds likewise vibrated their tails in response to sudden or rapid movements within their range of vision.

Taking precautions not to excite the rattlesnake, the stick with the cloacal secretion heavily smeared on the end was held in front of each of the sixteen rattlesnakes. None of them responded by assuming the "king-snake defense posture" even when the stick was held sufficiently close to the head for the tips of the tongue to touch it. If the stick was touched to the tip of the snout, the rattlesnake withdrew immediately as it had done when the clean stick came in contact with the snout.

To the stick that had been rubbed on the back of a king snake all rattlers responded by assuming the defense posture that had been evoked in the presence of the king snakes. Some responded more readily than others. The blindfolded *Crotalus s. scutulatus* responded before the tips of the tongue had come in actual contact with the stick. Six of the sidewinders, the *C. atrox*, the *C. v. lutosus*, and one *C. v. viridis* responded by assuming the "king-snake defense posture" as soon as their tongues touched the ends of the stick. One *C. v. viridis* responded only when a stick was applied to the neck simultaneously with the contact of the flickering tongue to the "scented" stick, but failed to respond if a stick was applied to the neck in the absence of scent. The remainder of the sixteen rattlesnakes were induced to respond only after repeated placings of the stick in proximity of the flickering tongue.

In the first series of tests with the stick that had been rubbed on the belly of the king snake the response was negative in all cases; that is, none of the rattlers assumed the "king-snake defense posture." When the same series of tests was repeated the following day nearly the same results were achieved as when the stick rubbed on the skin of the dorsum had been used, except that one *C. v. viridis* failed to respond. The evidence in the two series of tests using sticks rubbed

on the bellies of king snakes is considered faulty, either because of an imperfect technique in rubbing the stick prior to its use in the first series, or because of contamination by contact with skin on the dorsum in the second attempt. Unfortunately this part of the experiment could not be repeated owing to the desirability of completing the final test before any of the rattlesnakes became debilitated by prolonged starvation. Although the rattlesnakes drank water, all of them consistently refused the food that was offered with the exception of one sidewinder that devoured a lizard (*Dipsosaurus*).

A week elapsed before the final experiment, involving ablation of the tongue in half the specimens, could be undertaken. Twenty-four hours before tests were made, king snakes were again placed in the cages with each of the rattlesnakes in order to make certain that the crotalids to be used still responded to king snakes. Three *C. v. viridis* and one *C. cerastes* were now found to be indifferent to the presence of king snakes and were consequently omitted from subsequent tests. The Mojave rattlesnake (*C. s. scutulatus*) continued to react most violently, "slapping" at the king snake with its looped body when the *Lampropeltis* approached. Selecting only those rattlesnakes that continued to respond in the presence of king snakes, six *C. cerastes*, and one each of *C. v. viridis, C. v. lutosus, C. s. scutulatus,* and *C. atrox* were divided into two groups, one of which was to serve as the control group.

The entire trunk of the tongue was removed from five snakes, including *C. s. scutulatus, C. atrox,* and three *C. cerastes.* The *C. s. scutulatus* and one of the sidewinders remained blindfolded. It would have been desirable to cut the tongues off by degrees in order to repeat the experiment of Wilde (1938) who demonstrated that the tips of the tongue need not enter the organs of Jacobson in order for the feeding reaction to be evoked in garter snakes. However, this would have required additional time, and it was preferable to complete the experiment before the rattlesnakes were further weakened from lack of nourishment.

On each day for five days following the removal of the tongue, king snakes were placed in cages with the rattlers, both those with their tongues removed as well as those in the control group. None of the snakes without tongues assumed the "king-snake defense posture" after the operation, even though each of four king snakes was placed in the cage with them daily on five successive days for periods varying from five to thirty minutes. The rattlesnakes in the control group continued to assume the typical attitude, raising the vertical loop

in their bodies, and lowering their heads as they had done immediately after capture. Using both controls and snakes with the tongue removed tests were repeated ten days after the operation with similar results. At this time the evidence was considered satisfactory and no further tests were made. Snakes with their tongues ablated lived as long, on an average basis, as those with their tongues intact. The Mojave rattlesnake survived longer than any other of those used in the tests and did not succumb until approximately seventy-eight days after its tongue was removed.

REACTIONS OF RATTLESNAKES TO SERPENTS OTHER THAN KING SNAKES

Rattlesnakes have not been reported to respond to other species in the same way they respond to subspecies of *Lampropeltis getulus*. Despite the fact that several races of the species *getulus* are known to prey upon other snakes, the data for many other North American snakes are inadequate and too often based upon captive feedings. Many of the larger colubrids occasionally are ophiophagous, particularly those of the genus *Coluber* (Klauber 1927: 13; 1936: 206). Therefore, it was not surprising to find that rattlesnakes assumed the raised loop posture in the presence of certain colubrids in addition to king snakes.

Shortly after a juvenile *C. atrox* was captured near Indian Wells, California, a racer, *Coluber flagellum frenatum,* was placed in the cage with it. This rattlesnake was observed to assume the typical "king-snake defense posture" and similar observations were subsequently recorded for other rattlers in the presence of the racer. Although several other colubrids taken in the same region were subsequently placed in cages with *C. cerastes* and *C. atrox,* none of them evoked the characteristic response in the rattlesnakes. It is hardly worth while to name all the North American colubrids which failed to evoke any response in rattlesnakes, but a partial list is worth reporting. This includes several species of the genera *Thamnophis* and *Natrix, Pituophis catenifer annectens, Pituophis melanoleucus, Arizona elegans occidentalis, Rhinocheilus lecontei, Sonora occipitalis, Phyllorhynchus decurtatus perkinsi, Heterodon c. contortrix, Elaphe o. obsoleta,* and *Opheodrys vernalis.* Nor do rattlesnakes respond to other North American crotalids of the genera *Agkistrodon, Sistrurus,* or *Crotalus.*

Only two boids have been available, *Lichanura roseofusca gracia* and *Constrictor c. imperator.* Several species of rattlesnakes tested

were indifferent to both of these.   No elapid snakes have been tested, an unfortunate circumstance since many of them are ophiophagous.

In addition to the South American *Constrictor* mentioned above two live snakes of exotic origin were placed in cages with rattlesnakes. To one of these, the European *Natrix n. natrix*, rattlesnakes did not respond.   A South American species, *Pseustes sulphureus*, elicited the "king-snake defense posture" in all rattlers used in the tests described above.   Whether *P. sulphureus* is to any extent ophiophagous has not been ascertained.   It is perhaps noteworthy that *Crotalus durissus terrificus* occurs in British Guiana where the *Pseustes* was obtained.

Supplementing the series of tests performed with *Lampropeltis getulus californiae* of both pattern phases and other races of this species, one additional test was performed with the typical form of the eastern United States, *Lampropeltis g. getulus*.   Several species of rattlesnakes confined to the western portion of the continent, *C. atrox*, *C. s. scutulatus*, *C. v. viridis*, and *C. v. lutosus* reacted as readily to this king snake as to the western subspecies.   No species of *Lampropeltis* is known to occur in Montana where the *C. v. viridis* were taken.

Including observations recorded during the last ten years the following species and subspecies of rattlesnakes have been seen to assume the peculiar, characteristic defense posture in the presence of *Lampropeltis g. californiae:*

Crotalus enyo	*Crotalus v. lutosus*
Crotalus atrox	*Crotalus v. oreganus*
Crotalus ruber	*Crotalus mitchelii pyrrhus*
Crotalus s. scutulatus	*Crotalus cerastes*
Crotalus v. viridis	

As noted above a healthy *C. m. molossus* failed to react.   Whether this species is an exception, or whether the single specimen tested was merely an individual exception is not known.   In the past it has been noted that large specimens of *C. atrox*, even when tested immediately after capture, sometimes did not respond to the presence of the king snakes, although juveniles or young adults of the species responded without exception.   Since the *molossus* was a large specimen this may account for its failure to react.   Certainly most king snakes prey upon juvenile rattlesnakes or upon small species, and it may be found that a high percentage of large rattlesnakes do not react.

In addition to the species mentioned above, Meade (1940: 166) has reported the king-snake defense reaction of a canebrake rattlesnake

(*Crotalus horridus atricaudatus*) in the presence of the southern king snake (*Lampropeltis getulus holbrookii*).

SUMMARY

1. Normal healthy rattlesnakes of several species responded to the presence of ophiophagous king snakes, including four subspecies of *Lampropeltis getulus*, by assuming the characteristic defense posture described by Cowles. A similar response was evoked when a racer, *Coluber flagellum frenatum*, and a South-American colubrid, *Pseustes sulphureus*, were placed in the same cage with rattlesnakes. A rattlesnake secured from an area where no king snakes are known to occur, responded to the presence of a king snake. This response of rattlesnakes was not evoked by the presence of other crotalids, by boids, nor by several normally non-ophiophagous colubrid species used in the experiments.

2. Rattlesnakes assumed the "king-snake defense posture" when placed in a receptacle that previously had contained king snakes. This same reaction occurred when sticks that had been rubbed on the backs of king snakes were held immediately in front of rattlesnakes. Sticks upon which the secretion from the cloacal gland was smeared, evoked no response when held in front of rattlesnakes. Vibration of the rattle ceased when the reaction to king snakes was elicited in rattlesnakes, either by the presence of the king snake or by sticks previously rubbed on king snakes.

3. Blindfolded rattlesnakes responded as readily to the presence of king snakes or to sticks rubbed on the dorsal surface of king snakes as did rattlesnakes with normal vision.

4. After the initial stimulus from king snakes had evoked the characteristic response, rattlesnakes responded to visual or contact stimuli from other sources by assuming the posture previously assumed in the presence of king snakes. This response was evoked as long as three hours after the original stimulus was received *via* the tongue-Jacobson's organ mechanism.

5. Removal of the tongue in rattlesnakes, thereby eliminating an essential part of the tongue-Jacobson's organ mechanism, abolished the characteristic response.

DISCUSSION AND CONCLUSIONS

Whereas sex recognition has been investigated rather intensively by recent workers, particularly by Noble and Bradley (1933), by Davis (1936) and by Noble (1937), no experiments dealing with the

recognition of enemies have been reported beyond those of Cowles (1938). McIlhenny (1935: 44) asserts, apparently upon the basis of field observations, "Alligators seem to know the difference between poisonous snakes and non-poisonous snakes," and he describes the special method used by alligators in catching water moccasins (*Agkistrodon*). Similar casual observations have been made by other authors, but even such reports as these are scant.

The experiments described indicate that certain odorous substances present in the integument of king snakes as well as in certain other colubrids, serve as a stimulus which elicits a stereotyped response in rattlesnakes (*Crotalus*). This response, loosely termed a "defense posture," depends solely upon chemoreceptors, the organs of Jacobson. Immediately after the response has been evoked, visual and tactile stimuli may prolong it or cause it to be repeated, but the initial determining factor is chemical stimulation.

While exceptions have been noted, the response is of wide, if not of universal occurrence in snakes of the genus *Crotalus*. It is not known whether identical stimuli are present in the integuments of each of the three genera which evoked the response in rattlesnakes or whether rattlesnakes respond in a similar fashion to divers odors. Cowles' assertion that a similar response in rattlesnakes was evoked by the odor of a spotted skunk suggests that the latter interpretation may prove to be correct. Seemingly the reaction of rattlesnakes to other serpents is positively correlated with the ophiophagous propensities of the latter. However, more adequate data are required to establish this point satisfactorily. The response of the rattlesnakes tentatively may be interpreted as a "defensive activity," the origin and evolution of which would seem to be confined to speculative interest.

LITERATURE CITED

Baumann, F.
   1929. Experimente über den Geruchssinn und den Beuteerwerb der Viper (*Vipera aspis* L.). Zeit. vergl. Physiol. 10: 36–119. 36 text-figs.
Blanchard, F. N.
   1921. A revision of king snakes; genus *Lampropeltis*. Bull. U. S. Nat. Mus. No. 114. 1–260. Text-figs. 1–78.
Cowles, R. B.
   1938. Unusual defense postures assumed by rattlesnakes. Copeia 13–16.
Davis, D. D.
   1936. Courtship and mating behavior in snakes. Zool. Ser. Field Mus. Nat. Hist. 20: 257–290. Text-figs. 28–34.

Kahmann, H.
1932. Sinnesphysiologische Studien an Reptilien. I. Experimentelle Untersuchungen über das Jakobson'sche Organ der Eidechsen und Schlangen. Zool. Jahrb., Abt. Physiol. 51: 173–238. 31 text-figs.

Klauber, L. M.
1927. Some observations on the rattlesnakes of the extreme Southwest. Bull. Antivenin Inst. Amer. 1: 7–21.
1931. A statistical survey of the snakes of the southern border of California. Bull. Zool. Soc. San Diego. No. 8.
1936. The California king snake a case of pattern dimorphism. Herpetologica 1: 18–29.
1936. A key to the rattlesnakes with a summary of characteristics. Trans. San Diego Soc. Nat. Hist. 8: 185–276. Text-figs. 1–112.
1939. A further study of pattern dimorphism in the California king snake. Bull. Zool. Soc. San Diego. No. 15. 1–23. Text-figs. 1–7.

Manning, F. B.
1923. Hearing in rattlesnakes. Jour. Comp. Psychol. 3: 241–247.

McIlhenny, E. A.
1935. The alligator's life history. Boston: the Christopher Publishing House. 117 pp. 18 figs.

Meade, George P.
1940. Observations on Louisiana captive snakes. Copeia 165–168.

Mosauer, Walter
1930. A note on the sidewinding locomotion of snakes. Amer. Nat. 64: 179–183. Text-figs. 1–2.

Noble, G. K.
1937. The sense organ involved in the courtship of *Storeria, Thamnophis* and other snakes. Bull. Amer. Mus. Nat. Hist. 73: 673–725. Text-figs. 1–8.
————; & Bradley, H. T.
1933. The mating behavior of lizards; its bearing on the theory of sexual selection. Ann. New York Acad. Sci. 35: 25–100. 12 text-figs.
–; & Clausen, H. J.
1936. The aggregation behavior of *Storeria dekayi* and other snakes, with special reference to the sense organs involved. Ecol. Monog., 6: 269–316. 15 text-figs. 15 tables.

Wilde, Walter S.
1938. The role of Jacobson's organ in the feeding reaction of the common garter snake, *Thamnophis s. sirtalis* (Linn.). Jour. Exp. Zool. 77: 445–464. 1 pl.